

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ імені ІГОРЯ
СІКОРСЬКОГО»**

Факультет електроніки

(повна назва інституту/факультету)

Акустики та акустoeлектроніки

(повна назва кафедри)

«До захисту допущено»

Завідувач кафедри

Дідковський В.С.

(підпис)

(ініціали, прізвище)

“ ” 20__ р.

Дипломна робота
на здобуття ступеня бакалавра

зі спеціальності (спеціалізації) 6.050803 Акустотехніка
(код та назва спеціальності)

на тему: Моніторинг стану бджолиних сімей акустичним методом

Виконав: студент 4 курсу, групи ДГ-51
(шифр групи)

Шостак Станіслав Сергійович

(прізвище, ім'я, по батькові)

(підпис)

Керівник доцент, к.т.н. Заєць В.П

(посада, науковий ступінь, вчене звання, прізвище та ініціали)

(підпис)

Консультант

(назва розділу)

(посада, вчене звання, науковий ступінь, прізвище, ініціали)

(підпис)

Рецензент

(посада, науковий ступінь, вчене звання, науковий ступінь, прізвище та ініціали)

(підпис)

Засвідчую, що у цій дипломній роботі
немає запозичень з праць інших авторів
без відповідних посилань.

Студент _____
(підпис)

Київ – 2019 року

**Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»**

Інститут/факультет _____ електроніки
(повна назва)

Кафедра _____ акустики та акустoeлектроніки
(повна назва)

Рівень вищої освіти – перший (бакалаврський)

Спеціальність (спеціалізація) _____ 6.050803 Акустотехніка
(код і назва)

ЗАТВЕРДЖУЮ

Завідувач кафедри

_____ Дідковський В.С.
(підпис) (ініціали, прізвище)

« ____ » _____ 20__ р.

ЗАВДАННЯ

на дипломний проект (роботу) студенту

Шостак Станіславу Сергійовичу

(прізвище, ім'я, по батькові)

1. Тема проекту (роботи) Моніторинг стану бджолиних сімей акустичним методом

керівник проекту (роботи) _____ Заєць В.П., доцент, к.т.н.,
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом по університету від « ____ » _____ 20__ р. № _____

2. Строк подання студентом проекту (роботи) _____ 7 червня

3. Вихідні дані до проекту (роботи) база даних, створена на основі експериментальних досліджень, обробка отриманих даних в яку входить знаходження первинних класифікаційних ознак та оцінка їх інформативності, створення систем класифікації за класифікаційними ознаками та оцінка ймовірностей правильної класифікації за допомогою створених систем.

4. Зміст (дипломної роботи) пояснювальної записки (перелік завдань, які потрібно розробити) _____

1. Аналіз сучасного стану проблеми. _____

2. Організація експериментальних досліджень для збору даних. _____

3. Обробка отриманих даних. Результати досліджень. _____

5. Перелік графічного (ілюстративного) матеріалу (із зазначенням обов'язкових креслеників, плакатів, презентацій тощо) презентація

6. Консультанти розділів проекту (роботи)*

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв

7. Дата видачі завдання 5 вересня 2018 р

Календарний план

№ з/п	Назва етапів виконання дипломного проекту (роботи)	Строк виконання етапів проекту (роботи)	Примітка
1	Збір матеріалів для роботи. Аналіз науково-технічної літератури.	29.06.2017 - 10.12.2018	
2	Обробка отриманих даних в яку входить знаходження первинних класифікаційних ознак та оцінка їх інформативності.	10.12.18 - 10.02.19	
3	Створення систем класифікації конкретних станів бджолої сім'ї та оцінка ймовірностей правильної класифікації.	11.02.19 - 24.05.19	
4	Оформлення пояснювальної записки та презентації.	24.05.2019 - 06.06.2019	

Студент

(підпис)

С.С. Шостак
(ініціали, прізвище)

Керівник проекту (роботи)

(підпис)

В.П. Засць

* Консультантом не може бути зазначено керівника дипломного проекту (роботи)

РЕФЕРАТ

Моніторинг стану бджолиних сімей акустичним методом// Дипломна робота на здобуття ступеня вищої освіти «бакалавр». Шостак С.С. Національний технічний університет України «Київський політехнічний інститут ім. Ігоря Сікорського», факультет електроніки, кафедра акустики та акустoeлектроніки, група ДГ-51. – К.:НТУУ «КПІ», 2019. с. – 93, рис. – 70, табл. – 8.

Метою роботи є знаходження частотних складових звукового спектру бджолиної сім'ї, які дозволять надійно відрізнити такі стани бджолиної сім'ї як: стан колонії під час медозбору і його відсутності та стан колонії під час втрати королеви. Також класифікаційна задача включає в себе розробку правил прийняття відповідних рішень на основі отриманих інформаційних частотних складових.

Виконано експериментальні дослідження, спрямовані на виявлення ефективних класифікаційних ознак. Із використанням гребінки із одинадцяти 1/3-октавних фільтрів отримані оцінки спектральної потужності шуму бджіл в частотному діапазоні 125-1250Гц. Отримано оцінки інформативності значень спектральної потужності та виявлено, що для класифікації стану медозбору або його відсутності найбільш інформативними класифікаційними ознаками є оцінки спектру потужності шуму бджолиної сім'ї в околицях частот 200Гц та 250Гц. Для класифікації стану втрати королеви найбільш інформативними класифікаційними ознаками є оцінки спектру потужності шуму бджолиної сім'ї в околицях частот 800Гц та 1кГц. Створено дві системи класифікації для кожного з розглянутих станів. В першій системі класифікації в якості класифікаційних ознак використано два значення оцінок спектру на двох смугах інформативних частот, а в якості правила прийняття рішення – рівняння граничної кривої, що розділяє кластери конкуруючих класів на площі. В альтернативній системі класифікації в ролі класифікаційної ознаки використано оцінку спектру однієї з інформативних частотних смуг, а в якості

правила прийняття рішення –байєсівську стратегію. Тестування ефективності створених систем класифікації виконано шляхом оцінювання ймовірності правильної класифікації на навчальних та тестових вибірках даних. Таке тестування показало, що система класифікації що визначає наявність або відсутність медозбору, у випадку використання двох класифікаційних ознаки, забезпечує ймовірність правильної класифікації 0,81 для навчальної вибірки даних, та 0,96 для тестової вибірки даних. Аналогічне тестування було проведено для другого різновиду системи класифікації вищевказаного стану, де класифікаційною ознакою слугує лише одне значення спектру. Для випадку, коли це значення оцінювалося на частоті 200 Гц, ймовірність правильної класифікації склала 0,66 для навчальної вибірки даних та 0,90 для тестової вибірки даних. У випадку оцінки на частоті 250 Гц, ймовірність правильної класифікації становила 0,86 для навчальної вибірки даних та 0,84 для тестової вибірки даних.

Тестування ефективності системи класифікації, що визначає втрату бджолиної королеви у колонії, у випадку використання двох класифікаційних ознаки, забезпечує ймовірність правильної класифікації 0,87 для навчальної вибірки даних, та 0,84 для тестової вибірки даних. Аналогічне тестування було проведено для другого різновиду системи класифікації вищевказаного стану бджолиної сім'ї, де класифікаційною ознакою слугує лише одне значення спектру. Для випадку, коли це значення оцінювалося на частоті 800Гц, ймовірність правильної класифікації склала 0,78 для навчальної вибірки даних та 0,62 для тестової вибірки даних. У випадку частоти 1кГц, ймовірність правильної класифікації становила 0,88 для навчальної вибірки даних та 0,71 для тестової вибірки даних.

Ключові слова: бджолина сім'я, смуговий фільтр, поділяюча крива, класифікація, оцінка густини потужності.

ABSTRACT

Bee colonies condition monitoring by acoustic method // Thesis for a degree of higher education "Bachelor". S. Shostak, National Technical University of Ukraine "Kiev Polytechnic Institute named after Igor Sikorsky, Faculty of Electronics, Department of acoustics and acoustoelectronics, group DG-51. - K: NTUU "KPI", 2018. p. - 93, fig. - 70, tab. - 8.

The purpose of this research is to find frequency components of the sound spectrum of the bee colony, which will allow distinguishing such conditions of the bee colony as: the state of the colony during the honey harvest and its absence and the state of the colony during the loss of the queen. In addition, the classification task includes the development decision-making rules based on the informativeness frequency components.

Experimental researches were conducted with aim to identify effective classifications features. Using a comb from eleven 1/3-octave filters were obtained the spectral density estimates of bee noise in the frequency range 125-1250 Hz. The informativeness estimation of spectral density were obtained and was revealed that for the classification of honey gathering condition or its absence the most informative classification features are spectrum density estimation of the bee colony noise in frequency range 200 Hz and 250 Hz. For the classification of queen's loss condition the most informative classification features are spectrum density estimation of the bee colony noise in frequency range 800Hz and 1 kHz. Two classification systems for each of the considered conditions was implemented. In the first classification system, two values of the spectral estimates on two lanes of informational frequencies are used as classification features, and as a decision-making rule was chosen equation of the dividing curve, which used to separate the clusters of competing classes on the area. In an alternative system of classification, as a classification feature was used one of the informative frequency bands, and the Bayesian strategy was used as a decision-making rule. Testing of the classification systems efficiency was performed by assessing the probability of correct

classification in the training and test data samples. Such testing has shown that a classification system that determines the presence or absence of a honey gathering, in the case of using two classification features, provides a probability of a correct classification of 0.81 for the study sample data, and 0.96 for the test data sample. A similar test was conducted for the second variety of the classification system, where only one value of the spectrum density was considered as a classification features. For a case where this value was estimated at a frequency of 200 Hz, the probability of correct classification was 0.66 for the data sample training and 0.90 for the data sample test. In the case of a frequency of 250 Hz, the probability of correct classification was 0.86 for the data sample training and 0.84 for the test data sample.

The testing of the efficiency of the classification system, which determines the loss of a queen bee in the colony, in the case of using two classification features, provides a probability of a correct classification of 0.87 for the training data sample, and 0.84 for the test data sample. A similar test was conducted for the second variety of the classification system of the mentioned condition of the bee colony, where only one value of the spectrum density was considered as a classification feature. For the case when this value was located at 800Hz, the probability of correct classification was 0.78 for the training data sample and 0.62 for the test data sample. In the case of a frequency of 1 kHz, the probability of correct classification was 0.88 for the data sample training and 0.71 for the data sample test.

Key words: bee colony; band-pass filter; dividing curve; classification; power spectrum density.

ЗМІСТ

ВСТУП	10
РОЗДІЛ 1. АНАЛІЗ НАУКОВО-ТЕХНІЧНОЇ ЛІТЕРАТУРИ.....	12
1.1. Встановлення зв'язку між шумом бджолоїної сім'ї та її біологічним станом.....	12
1.2. Застосування методів спектрального аналізу для знаходження стійких класифікаційних ознак	17
1.3. Виявлення статистично однорідних ділянок шуму бджолоїної сім'ї для діагностики стану колонії	29
1.4. Використання алгоритмів машинного навчання для моніторингу бджолоїних сімей.....	33
РОЗДІЛ 2. ОРГАНІЗАЦІЯ ЕКСПЕРИМЕНТАЛЬНИХ ДОСЛІДЖЕНЬ ДЛЯ ЗБОРУ ДАНИХ. ПОПЕРЕДНЯ ОБРОБКА ДАНИХ.	42
2.1 Загальний опис та мета проведених досліджень	42
2.2. Організація першого експериментального дослідження.....	42
2.3. Попередня обробка даних з першого експериментального дослідження	45
2.4. Організація другого експериментального дослідження	49
2.5. Попередня обробка даних з другого експериментального дослідження	52
2.6. Організація третього експериментального дослідження.....	56
2.7. Попередня обробка даних з третього експериментального дослідження	60
2.8. Організація четвертого експериментального дослідження.....	61
2.9. Попередня обробка даних з четвертого експериментального дослідження	63

РОЗДІЛ 3. ОБРОБКА ОТРИМАНИХ ДАНИХ. РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ	64
3.1. Обробка даних, отриманих у результаті першого експерименту...	64
3.2. Обробка даних, отриманих у результаті другого експерименту. ..	69
3.3. Оцінка інформативності первинних класифікаційних ознак для даних, отриманих у результаті четвертого експерименту	74
3.4. Створення систем класифікації для контролю активності бджіл під час медозбору.	75
3.5. Порівняння ефективності систем класифікації для контролю активності бджіл під час медозбору	81
3.6. Оцінка інформативності первинних класифікаційних ознак для даних, отриманих у результаті третього експерименту	82
3.7. Створення систем класифікації для виявлення втрати королеви у колонії	83
3.8. Порівняння ефективності систем класифікації для виявлення втрати бджолої королеви у колонії	87
ВИСНОВКИ.....	89
СПИСОК ЛІТЕРАТУРИ.....	92

ВСТУП

Бджільництво є однією з найдавніших форм сільського господарства тому ця дисципліна поступово прогресувала разом з людськими знаннями і технологіями. Цей напрямок сільського господарства є надзвичайно важливим, оскільки дозволяє отримати низку продуктів, унікальних за хімічним складом, зокрема таких як мед, прополіс і маточне молочко. Бджоли є надзвичайно цінними комахами ще й тому, що опилюють рослини, підтримуючи таким чином екологічний баланс. Підраховано, що медоносні бджоли відповідають за запилення понад 90% послуг глобального комерційного запилення, та приблизно 35% світових продовольчих культур [1]. Технологічний прогрес ставить нове завдання для даної галузі: моніторинг стану бджолиних сімей, що дозволить здійснювати віддалений контроль за ними. Виконання завдання по оперативному й точному визначенню стану бджолиної сім'ї акустичним методом є актуальним, оскільки дозволяє поліпшити виробництво якісних продуктів бджільництва, зменшити втрати продуктів бджільництва та ресурсів бджоляра.

Для поліпшення виробництва і збуту продукції бджільництва, а також для підтримки екосистеми повинні бути складені системи класифікації що забезпечать виявлення типових проблем у виробництві, а саме: втрата королеви, роїння [2], контроль за активністю бджіл під час медозбору, наявністю захворювання на вроатоз [3]. На основі вищезгаданих систем класифікації повинна бути створена єдина система акустичного моніторингу бджолиної сім'ї. Розв'язання такого завдання складається із трьох етапів. На першому етапі створюється база даних, що складається з аудіо записів шуму бджолиної сім'ї у різних класифікаційних станах. Отримана база даних ділиться на навчальну та тестову вибірку. На другому етапі, шляхом аналізу взаємозв'язків між класифікаційним станом бджолиної сім'ї та характеристиками шуму, який вона призводить, виконується вибір класифікаційних ознак. На третьому етапі будується система класифікації

стану сім'ї за обраними класифікаційними ознаками, та оцінюється ефективність даної системи на навчальній та тестовій вибірці. В рамках даної атестаційної бакалаврської роботи представлено виконання всіх зазначених етапів та створено дві системи класифікації. Одна з систем дозволяє виявити втрату бджолої королеви. Друга – дозволяє відслідковувати активність бджіл під час медозбору.

РОЗДІЛ 1. АНАЛІЗ СУЧАСНОГО СТАНУ ПРОБЛЕМИ

1.1. Встановлення зв'язку між шумом бджолої сім'ї та її біологічним станом

Одну із перших спроб встановити зв'язок між характером звуків, що генерує бджолої сім'я, та біологічним станом бджіл за допомогою методів акустичного аналізу було виконано Е.Ф. Вудсом [4]. Автор представив аналіз звукового спектру великої здорової колонії бджіл у трьох умовах на рисунку 1. Кожна з умов показана на кривих I, II і III, що являють собою обвідні звукового спектру. Ці криві, отримані двома методами: шляхом відтворення потрібного звуку за допомогою фільтрувального ланцюга і прив'язки його до наближеного аналогічного тону, визначеного на слух; і використання спектроаналізатора. Ці два методи дали схожі результати, а другий був прийнятий за основний, будучи набагато простішим і швидшим.

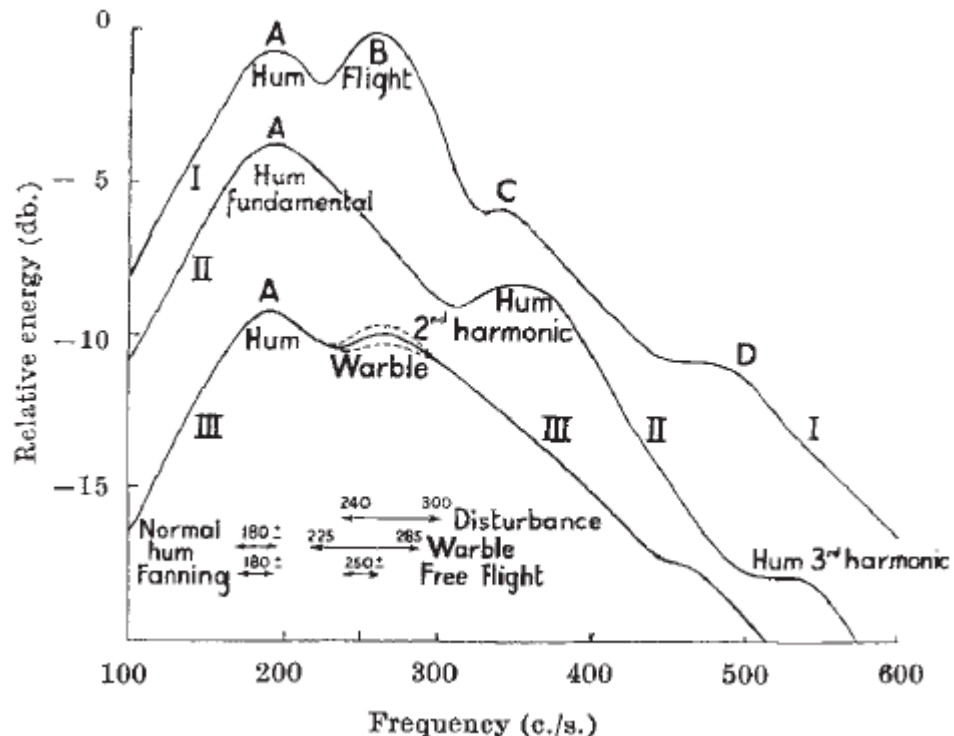


Рис.1. Обвідні спектру для трьох різних умов.

Крива I показує обвідну спектру під час звичайного робочого дня колонії.

Автором описано складові звуків бджолоїної сім'ї в діапазоні від 180 Гц до 3000 Гц, при цьому виділено чотири основні складові, що підписані на рисунку 1 як А,В,С і D. Складова А є основною, вказує на частоту 180 ± 5 Гц та породжується дихальними процесами бджіл. Складова В (250 ± 1 Гц) пов'язана зі звуками польоту та з настанням темряви зникає. С і D являють собою гармоніки А і В що зливаються на високих частотах. Четверта складова займає область частот близько 3000 Гц та генерується крилами бджіл в період підготовки бджіл до роїння, а також в період зимування бджіл.

Крива II показує обвідну спектру тієї ж колонії після настання темряви. Рівень основних частот спадає, та зникає складова В.

Крива III показує обвідну спектру тієї ж колонії при відсутності літальної активності бджіл та при вилученні бджолоїної королеви з сім'ї. Записи були зроблені через 3 години після того як матку було вилучено. В цьому випадку було виявлено інформаційні частоти (від 225 до 285 Гц), та їх флуктуації відносно рівня на ± 3 дБ. Автор пов'язує виявлені частоти саме з вилученням матки. Також автор вважає що сигнали на цих частотах випромінюється бджолами-годувальницями у віці від 4,5 до 6 днів в період підготовки сім'ї до роїння і протягом п'яти годин після вилучення з вулика матки.

У своєму дослідженні Дж. Джиндра [5] вказує, що бджоли створюють звук між 20 і 10 кГц, але найбільш інформативна частина спектру знаходиться в смузі від 200 до 500 Гц. Він описує, що бджолоїні колонії виробляють більш високу інтенсивність звуку при 240 Гц до 21 дня перед роїнням. Цей стан означає що бджоли почали створювати більше десятка нових королев. За 8 днів до рою максимальна інтенсивність звуку рухається до 270 Гц. За 1 день до рою максимальна інтенсивність зареєстрована на частоті 300 Гц. Автор також зауважує, що колонія бджіл виробляє більш високу інтенсивність в діапазоні 300 – 400 Гц, на 4 – 5 годину після втрати королеви.

Д.Г. Дейтлейн [6] у своїх роботах досліджував розподіл рівня амплітуди рівня фонового шуму вулика у денному циклі впродовж всіх чотирьох сезонів. Типовий приклад рівня фонового шуму вулика в добре розвинутій колонії за 5-денний період (26 лютого – 1 березня 1980) показаний на рисунку 2. У більшості днів був щоденний ритм мав підвищення амплітуди близько 8 години і поступово знижувався після настання темряви.

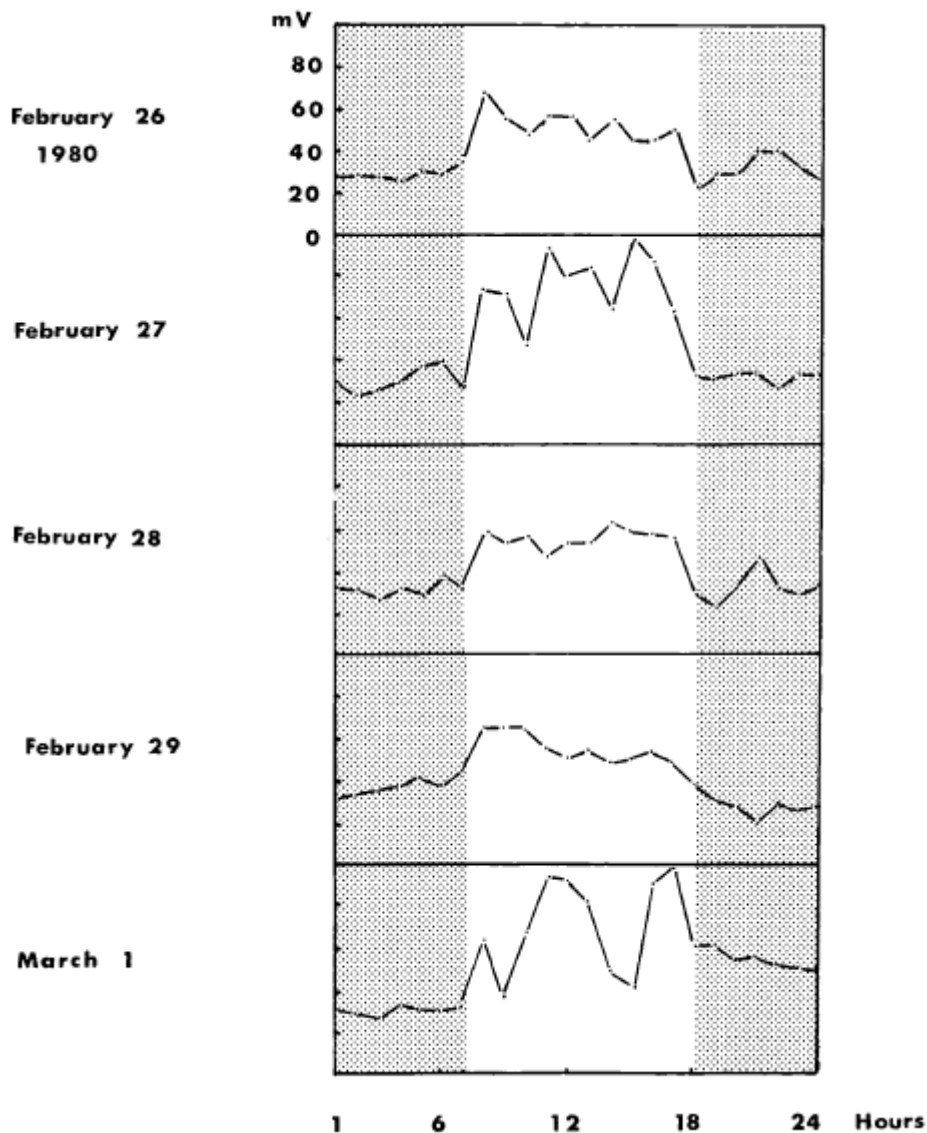


Рис.2. Погодинний розподіл амплітуди фонового шуму в денному циклі з лютого по березень.

Типова щоденна закономірність для розвинутої колонії навесні наведена на рисунку 3. Вона є подібною до попередньо розглянутої, за винятком значно

більш тривалого інтервалу між сходом сонця і збільшенням рівня звуку. Рівень фонового шуму записувався інтервалом в одну годину протягом дня і ночі, восени та влітку. Запис загального рівня звуку в травні показав чіткий щоденний цикл з часом підйому між 8 і 10 годинами (рисунок 3). Зниження рівня звуку відбувалося близько півночі.

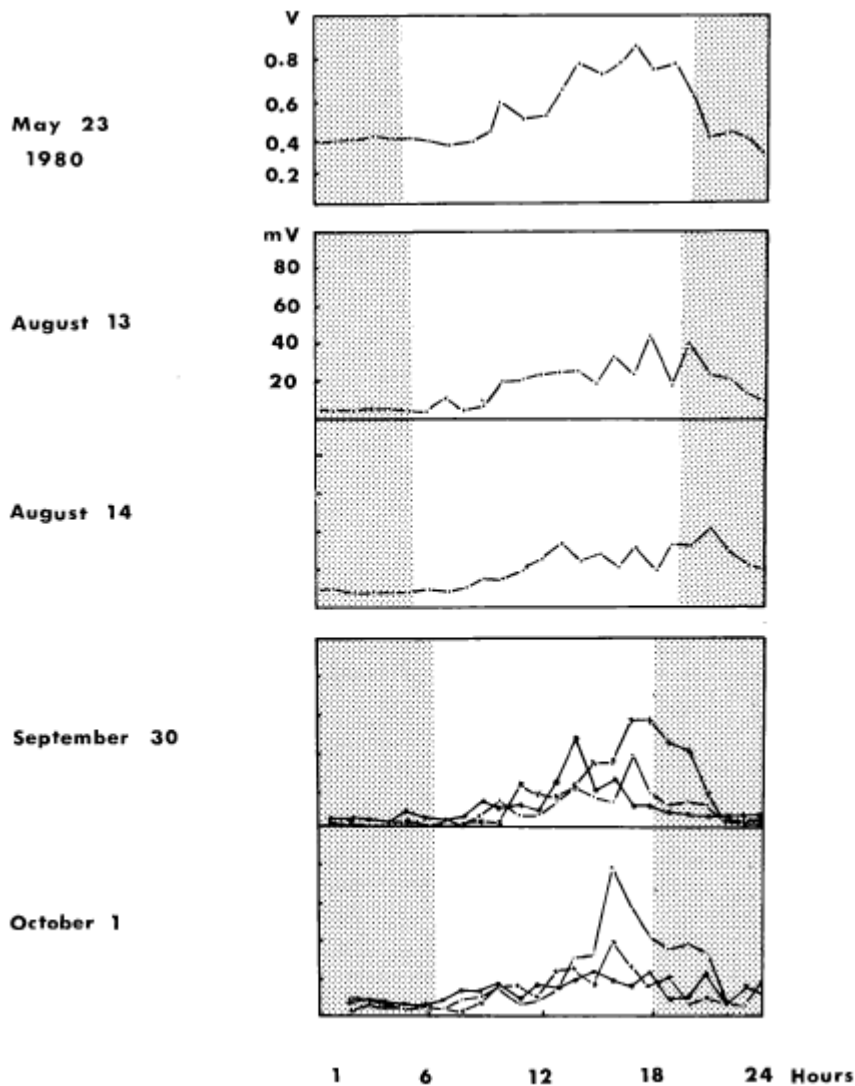


Рис.3. Погодинний розподіл амплітуди фонового шуму в денному циклі з кінця березня по початок жовтня.

Схема літньої активності була схожою до весняної, у тому числі при відставанні на 3-4 години між сходом сонця і початком збільшення активності колоній. У теплу погоду, коли рясні джерела нектару були присутні, рівень фонового шуму продовжувався навіть після опівночі.

Рівень шуму колоній восени (на 30 вересня і 1 жовтня) було також зображено на рисунку 3, для ілюстрації подібності по амплітуді при денному розподілі, в цьому сезоні. З настанням осені відбувалося зменшення пікової амплітуди до тих пір, поки рівні звуку не стали рівномірними, дуже слабкими і не мали періодичності.

Потрібно було підсилювати сигнал вдвічі, ніж влітку, щоб показати осінні амплітудні тенденції чітко. Рівні звуку мали тенденцію до зниження амплітуди, коли температура повітря ставала холоднішою, і висота основного піку також зменшувалася, поки до грудня його вже не можна було виділити. Це зображено на рисунку 4.

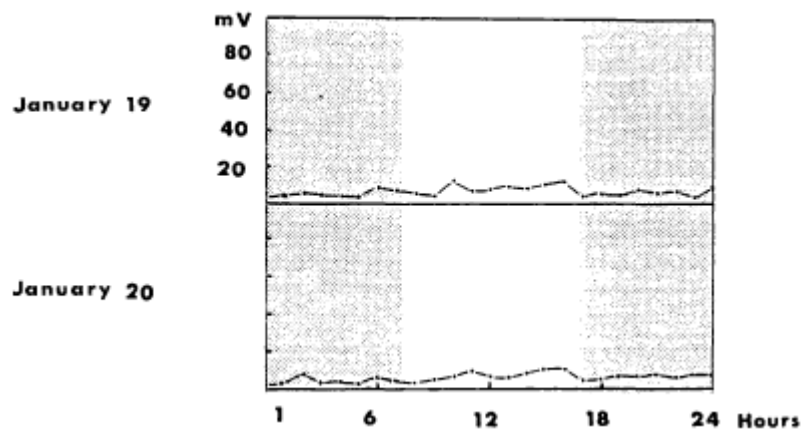


Рис.4. Погодинний розподіл амплітуди фонового шуму в денному циклі взимку.

Що стосується спектрального аналізу – автором було вказано що

більшість звуків вулика мали частоту в 300, 410 і 510Гц згідно фільтрів аналізатора спектра, з невеликим звуком вище 1кГц. Спектр змінювався з сезоном. Спектри в зимуючих колоніях були подібними в грудні і на початку січня, але відрізнялися у другий тиждень січня, коли частоти нижче 300Гц були повністю виключені зі спектру колоній.

1.2. Застосування методів спектрального аналізу для знаходження стійких класифікаційних ознак

В працях Е.К. Еськова [7], [8], була досліджена динаміка структури звуків бджолої сім'ї в річному і добовому циклі, а також зв'язки між змінами в цих структурах і біологічним станом бджіл. Автор вказує на існування певних закономірностей, що породжують складові звукового фону вулика. В процесі цих досліджень визначена спектральна і спектрально-часова структура звукового фону вулика. Зокрема, автором було встановлено, що спектр сім'ї, яка перебуває в нормальному стані, укладається в діапазоні від 20 до 8000-12 000Гц (динамічний діапазон інтенсивності 40дБ). Максимум енергії фону зосереджується в частотних смугах 135-190, 210-330 і 400-550Гц. У більшості випадків, за даними, Еськова максимально інтенсивні складові фону зосереджуються в діапазоні 210-330 Гц. Цей стабільний частотний діапазон вдень і вночі, взимку і влітку присутній на звуковому тлі вулика. Діапазон частотних складових звукового фону вулика (210-330Гц) характеризується тим, що частотне положення піку інтенсивності в ньому визначається біологічним станом бджіл. На частотну локалізацію піку інтенсивності в цьому діапазоні істотно впливає рівень активності бджіл під час медозбору, з підвищенням якого пік інтенсивності зміщується на високочастотну область частотного діапазону. Цей зсув носить ступінчастий характер, причому в період активізації бджіл пік інтенсивності зміщується на більшу величину, ніж після стабілізації рівня медозбиральної активності.

Поява в сім'ї групи малоактивних бджіл перед ройовим періодом відображена через посилення інтенсивності в низькочастотній області. З наближенням дня виходу рою інтенсивність низькочастотних складових (210-240Гц) частотного діапазону збільшується. За 5-15 днів до виходу рою співвідношення низькочастотної і пікової областей частотного діапазону змінюється по відношенню до вихідного рівня на 8-12 дБ. За цим показником звукового фону вулика сім'ї, що готуються роїтися, можна чітко відрізнити від

сімей, розвиток яких проходить минаючи роїння і сімей, які вийшли з ройової фази.

Автором було зазначено що підвищена активність ройових бджіл в перші 18-20 днів після заселення їх в нове житло відображається локалізацією піку інтенсивності в високочастотній області. Зміна локалізації піку інтенсивності в стабільному діапазоні частот можна викликати впливом зовнішніх стимулів, що впливають на зміну активності бджіл. Наприклад, на 20-30Гц зміщується пік інтенсивності на область високих частот в результаті рясного годування бджіл цукровим сиропом, розташованим в вулику або за його межами. При цьому відбувається посилення фону складових стабільного частотного діапазону на 12-15дБ. Підвищення активності зимуючих бджіл відображається на посиленні видаваних ними звуків, зміщенні піку інтенсивності в стабільному діапазоні на область високих частот і підвищення тепловиділення.

Автором було також вказано що зміна мікроклімату у вулику впливає на інтенсивність складових в діапазоні 135-190Гц. Найбільшої інтенсивності цей діапазон досягає в літній період при високій температурі, коли в вулик вноситься велика кількість нектару. Різко зростає інтенсивність цього діапазону в період зимівлі при підвищенні в вулику температури і концентрації вуглекислого газу. У цей період при підвищенні температури з 1-2° до 10-13° інтенсивність складових в діапазоні 135-190Гц зростає по відношенню до складових стабільного діапазону на 18-20дБ. При підвищенні вмісту в вулику вуглекислого газу з 2 до 5-7% (зимовий період) інтенсивність складових в області 135-190Гц зростає по відношенню до інтенсивності частотного діапазону на 5-6дБ. Якщо подавати в вулик атмосферне повітря, інтенсивність цього діапазону зменшується в міру зниження концентрації вуглекислого газу.

Далі автор стверджує що на основі наявних відомостей можна віднести складові в діапазоні 210-330Гц до процесів що пов'язані з тепловиділенням

бджіл. Цей висновок ґрунтується на наступному: по-перше цей діапазон постійно присутній в звуковому тлі вулика, а бджоли, як відомо, в результаті постійного тепловиділення підтримують в гнізді певний тепловий режим. По-друге зниження температури в периферійній зоні гнізда відбивається на зменшенні інтенсивності складових даного частотного діапазону. По-третє бджоли здатні виробляти звук в діапазоні цих частот виключно вібрацією тораксу під дією крилової мускулатури, в процесі чого виділяється тепло.

Звуки в діапазоні 135-190Гц автор пов'язує з процесами вентилявання. Про це говорить їх зв'язок з активністю вентиляції в бджолиної сім'ї: вони посилюються при необхідності зниження температури і вмісту вуглекислого газу.

Складові в діапазоні 400-550Гц, за словами автора, породжуються, в основному, сигналами, що мають мобілізаційне значення для мешканців вулика. Інтенсивність цього діапазону зростає перед виходом і в момент виходу рою, а також в період обльотів. Сильне збудження бджолиної сім'ї під впливом стимуляції, різко змінює функціональний стан бджіл, істотно впливає на посилення складових в діапазоні 400-550Гц.

Таким чином, автор зауважує, що по спектральній структурі звукового фону вулика можна судити не тільки про біологічний стан бджіл (по співвідношенню інтенсивності складових в діапазоні 210-330Гц), а й про те, наскільки умови житла відповідають біологічним вимогам сім'ї (за складовими в діапазоні 135-190 Гц).

В своїй науковій роботі С. Феррарі, Д. Сілва і М. Гуаріно [9] провели експеримент для ідентифікації роїння італійської медової бджоли *Apis Mellifera Ligustica* за допомогою методів класифікації. З трьох вуликів були зібрані звукозаписи та показники вологості і температури всередині вуликів. Запис проходив безперервно протягом 270 годин в травні 2006 року на експериментальній фермі Міланського університету. Для збору якомога більшої кількості записів під час роїння бджоляр не видаляв новозбудованих

маточників. Запис проводився з трьох мікрофонів (ЕСМ 3005, Monacor) з частотною характеристикою від 50 Гц до 16000 Гц. Записаний сигнал був оцифрований звуковою картою Soundscape SS88IO-3 з 8 каналами та квантуванням в 16 біт. Також використовувалися попередні підсилювачі, які дозволяли регулювання рівня вхідного сигналу та уникання перебільшення діапазону амплітуди сигналу що транслюється в АЦП.

Для контролю температури і відносної вологості дослідниками були використані реєстратори даних НОВО (Датчик Н08-007-02). Передача даних відбувався кожні 3 хв. Обидва датчика (мікрофон і НОВО) були захищені від прополісування спеціальною сіткою.

Операції обробки сигналів були зроблені за допомогою програмного забезпечення Matlab 7.1.

Першим етапом аналізу в даному дослідженні було фільтрування звукового сигналу щоб позбутися від навколишнього шуму, та відокремити частоти які нас цікавили. А саме, фільтр Баттерворта [10] (смуговий фільтр 6-го порядку) з частотами відсікання 100 Гц і 2000 Гц. Після фільтрації звукові сигнали і дані отримані з датчиків були синхронізовані. Під час запису відбулося дев'ять випадків ройових станів, які були виявлені як за акустичним моніторингом, так і візуально, спостерігачем. Виліт рою відбувалися завжди в денний час і переважно у найбільш спекотні години.

Аналізуючи записаний звук, було отримано інформацію про діяльність бджіл особливо під час роїння та під час нормальної діяльності колонії вдень та вночі. Ці варіації показані на рисунку 5, у вигляді спектрограм звукових записів протягом ночі, дня і періоду роїння.

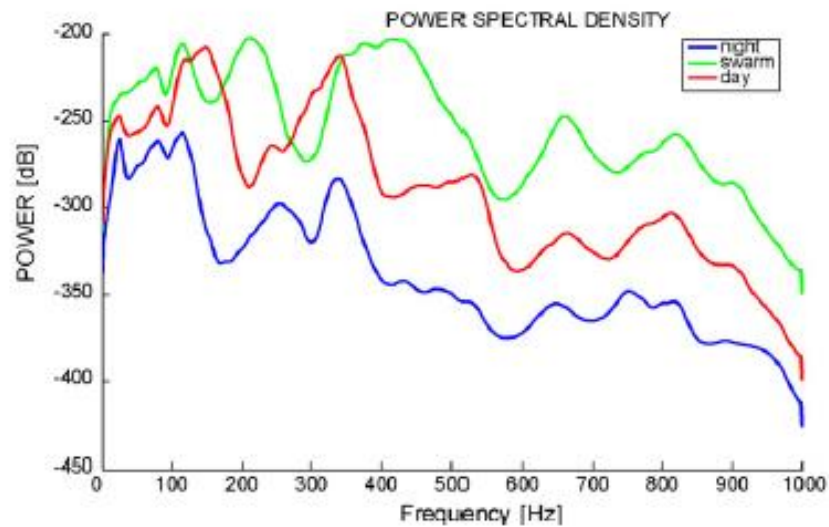


Рис.5. Спектрограми звукових записів протягом ночі (крива синього кольору), дня (червоного кольору) і періоду роїння (зеленого кольору).

У нічний час загальна амплітуда набагато нижче у порівнянні зі звуком під час нормальної денної активності. З рисунку видно, що найвищі показники інтенсивності знаходяться приблизно на однакових частотних смугах. При порівнянні спектрограм під час роїння з іншими то основною відмінністю є зсув на частотних смугах 150-250Гц і 350–450Гц в напрямку високих частот. Також спостерігається приріст по інтенсивності на високих частотах в порівнянні з нормальною активністю бджіл.

Зібрані дані свідчать про чітке зростання інтенсивності коли колонія знаходиться в процесі роїння в порівнянні з іншими ситуаціями.

На рисунку 6 показано перехід бджіл до ройового стану.

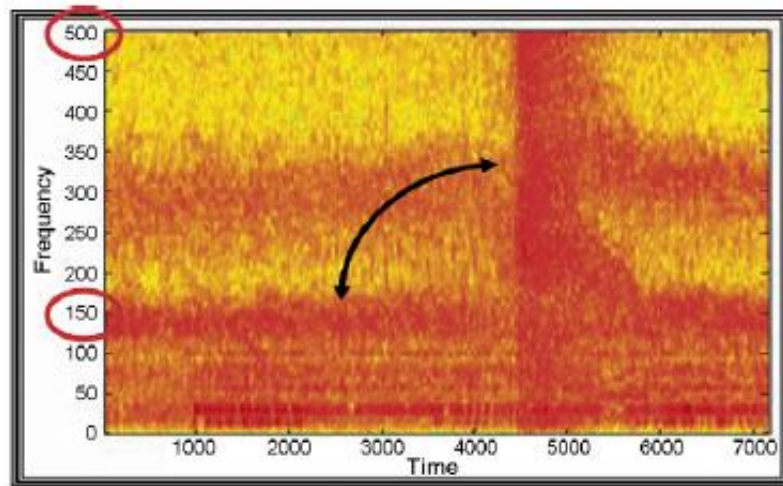


Рис.6. Спектрограма звукозаписів колонії у ройовому стані.

Можна побачити, що домінуючі смуги частот (50–150Гц і 250–350Гц) розширюються і відбувається збільшення інтенсивності на всьому частотному діапазоні. На спектрограмах можна виділити піки 25-150 Гц для активності в нічний час, піки 100-150–300Гц в денний час і піки в період 500–600Гц під роїння. Отже, дослідник довели що частотна область може бути якісним інструментом для оцінки наближення до певного явища оскільки швидке зростання піків відбувається в області високих частот при роїннях і дозволяє зробити чітке розмежування звуку при ройовому стані колонії та звуку при нормальній денній активності. Також існує зв'язок між спектральними характеристиками і температурою, що змінюється протягом дня. За декілька хвилин до вильоту відбувається зниження температури від 35 °С до 33 °С у фактичний час вильоту рою температура знову падає до 32°С, це показано на рисунку 7.

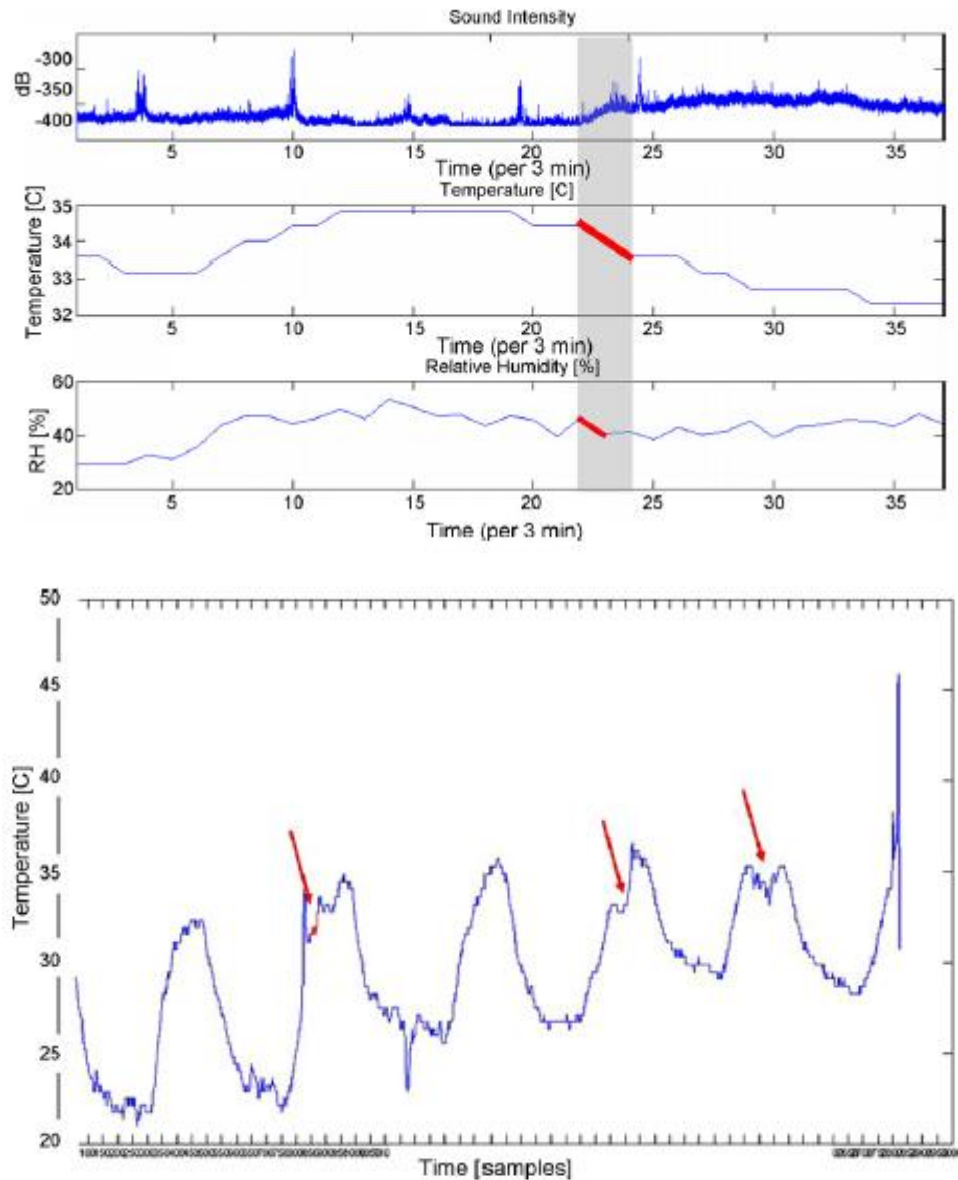


Рис.7. Розподіл температури в колонії під час ройового стану.

Можливим поясненням цієї зміни автори вважають явище вентиляції, яке полягає в швидкому руху крил бджіл для досягнення температури близької до 32°C . Перед роїнням ми також спостерігаємо зменшення відносної вологості що також ймовірно обумовлено феноменом вентиляції.

Д. Атаурі - Мезкіда та Дж. Лорент - Мартінез [11] розробили систему для акустичного моніторингу бджолиних сімей. Структурна схема цієї системи зображена на рисунку 8.

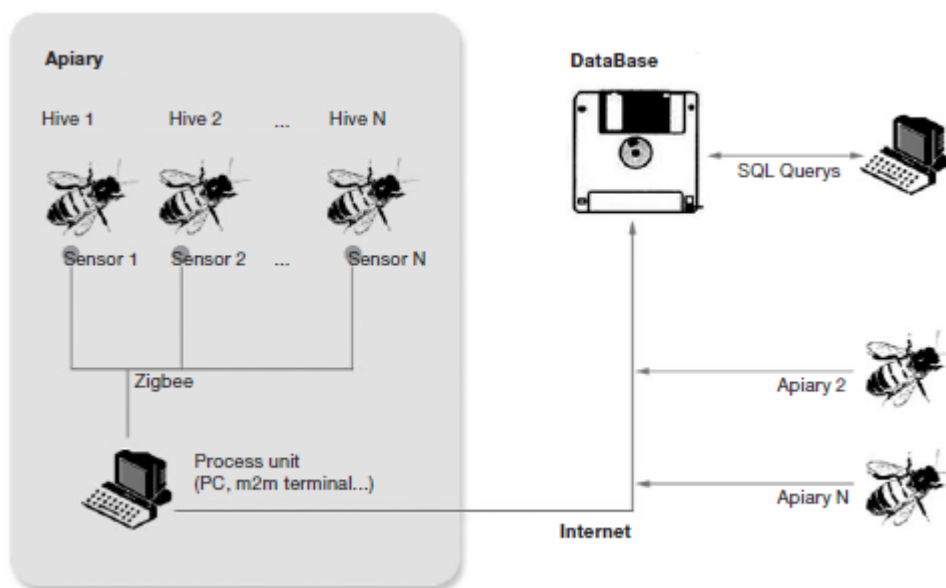


Рис.8. Структурна схема системи збору, обробки та збереження даних, використаної в цьому дослідженні.

Дані у вигляді звукозапису отримується з вибраного сенсора, потім обробляються і відправляються в базу даних. Сенсори були спроектовані і побудовані з наступних компонентів: всенаправлений мікрофон, підсилювач, АЦП з розрядністю в 8 біт, датчик температури, датчик бездротового зв'язку Zigbee, цифровий процесор сигналів. Авторами було розроблено програмне забезпечення для збереження звуку та його обробки. Інтернет з'єднання, необхідне для заповнення бази даних, може бути виконане через 3G модем або Wi-Fi. Щогодини з 10 вуликів записувались зразки шуму всередині колонії. Кожен звукозапис триває 8 секунд. Частота дискретизації встановлена на 6250Гц, так що найвища частота яку може прийняти дана система становить 3125Гц, відповідно до теореми Найквіста. Щоб отримати їх частотний спектр звукозаписів та визначити основні частотні складові було використано перетворення Фур'є. Отримані спектрограми зображені на рисунку 9.

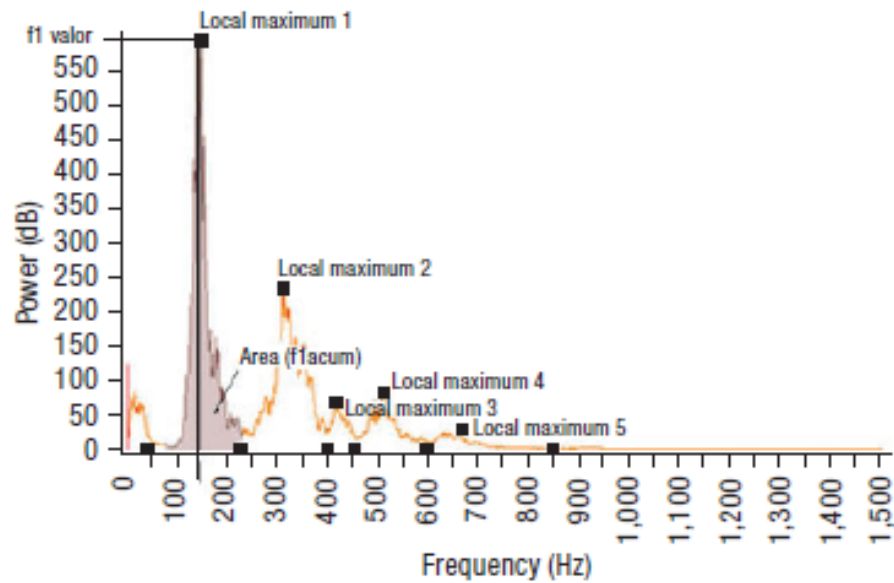


Рис.9. Спектрограма типового зразка звукозапису.

Всі дані отримані з різних звукових зразків зберігалися в базі даних, тому їх можна було легко фільтрувати, сортувати та групувати за різними критеріями за допомогою мови програмування SQL. Записи почалися 20 травня 2008 року до 4 квітня 2009 року було збережено 26774 звукозаписів.

База даних була реалізована в MySQL 5.0. Програми для обробки сигналів були розроблені за допомогою Borland C++ Builder 6. Програмне забезпечення сенсорів розроблено за допомогою ANSI C. Автори вказують що інтенсивність звуку на середніх і низьких діапазонах частот показала щоденні закономірності, насамперед восени. На рисунку 10 показано, як змінювалась інтенсивність сигналу в залежності від часу на трьох різних вуликах ранньою весною.

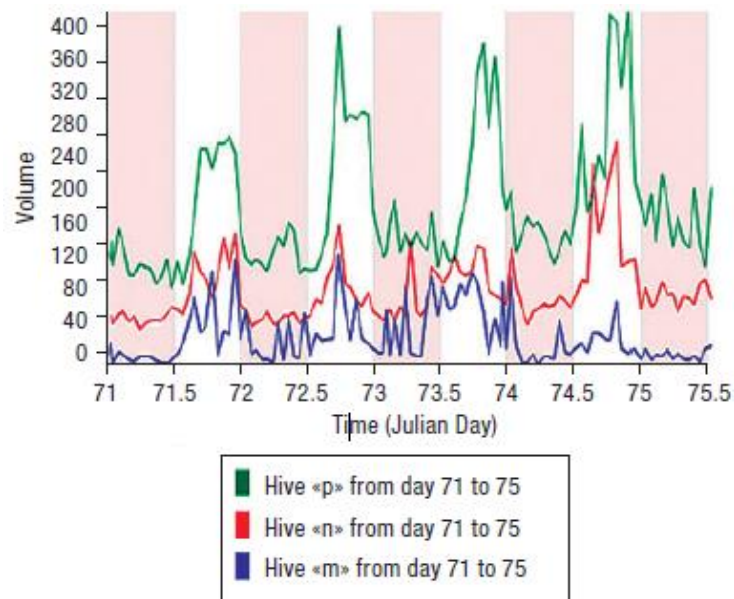


Рис.10. Варіації гучності трьох різних вуликів.

Зелена лінія відповідає вулику з вісьмома рамками розплоду, червона лінія відповідає вулику з п'ятьма рамками розплоду у вулика; синя лінія – трьом рамкам розплоду які були інфіковані *Ascosphaera apis*. На рисунку 10 зображено важливе співвідношення між гучністю звуку та кількістю розплоду у сім'ї. Інтенсивність звуку була зазвичай низькою на високих частотах (вище 300Гц), але значно підвищувалася при стресі, на зразок після приєднання бджіл з інших сімей. Деякі інші параметри, такі як основний тон і інформативні смуги частот, були в основному постійними та але за певних обставин показали різкі зміни. Спектральні характеристики зображені на рисунку 11.

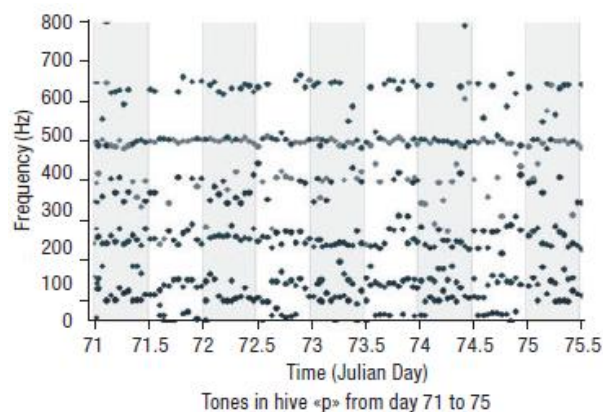


Рис.11. Варіації тону на основних смугах частот в одному вулику.

У даному дослідженні автори не пов'язують значення отриманих параметрів з конкретними станами бджолиних сімей, але зазначають що в подальшому можливо розкрити ці зв'язки.

М. Бенсік, Дж. Бенсік і М. Бакстер [12] у своїй дослідницькій роботі зазначали що вимірювані коливання походять від декількох тисяч медоносних бджіл, які можуть розглядатися як сукупність перехідних імпульсних осциляторів з випадковими окремими фазами. Отриманий сигнал, таким чином, вважали занадто шумним для прямого аналізу.

Обчислюючи усереднені вузько-частотні спектри (eigenspectra), була отримана чітка крива з вираженими піками, що свідчить про те, що коливання, викликані осциляторами, мають подібні спектральні характеристики, незважаючи на відсутність когерентності фаз. Тривалість часу між окремими відділками для обчислення спектру визначає його роздільну здатність за частотою, тоді як тривалість часу, протягом якої вона усереднюється, визначає відношення сигнал – шум сигналу. Дослідниками було знайдено компроміс для роздільної здатності частоти 20 Гц і часу усереднення 510 секунд. Використовуючи Matlab спектри що були зібрані за один день були усереднені для всього набору даних і були складені спектрограми для візуального контролю. Дані були отримані з двох вуликів. Для знаходження інформаційних ознак було використано аналіз основних компонентів (principal component analysis – PCA). Залежність частоти від амплітуди за денний цикл для двох вуликів показані на рисунку 12.

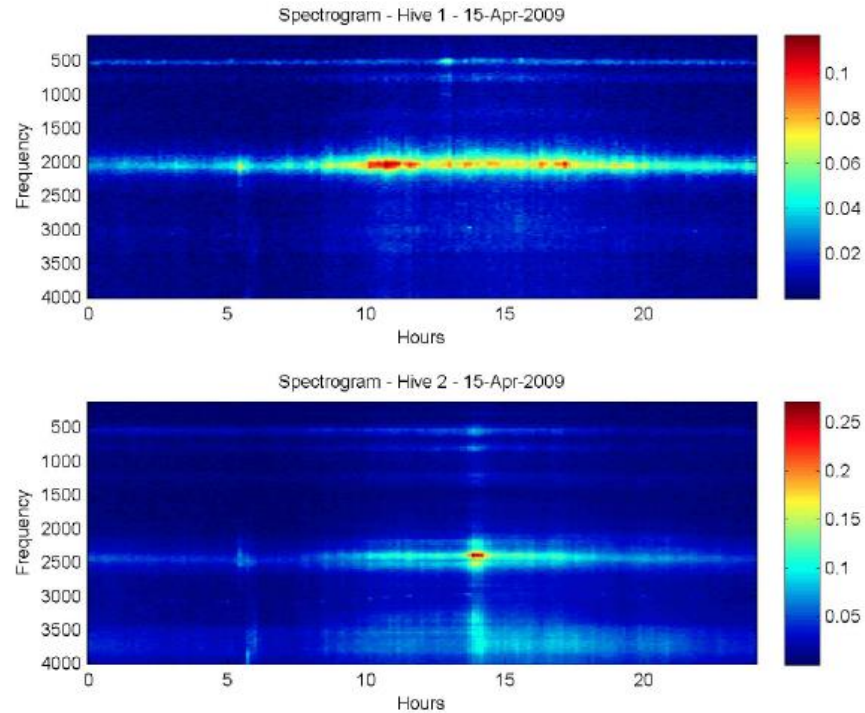


Рис.12. Залежність частоти від амплітуди за денний цикл.

Для обох вуликів чітко видно від чотирьох до п'яти «смуг», а найбільш виражені - близько 2000Гц. Відмінності між частотними характеристиками двох вуликів незначні. Амплітуда сигналу в другому вулику менша ніж в першому. Усереднені спектри, які були отримані в цьому дослідженні, показані на рисунку 13.

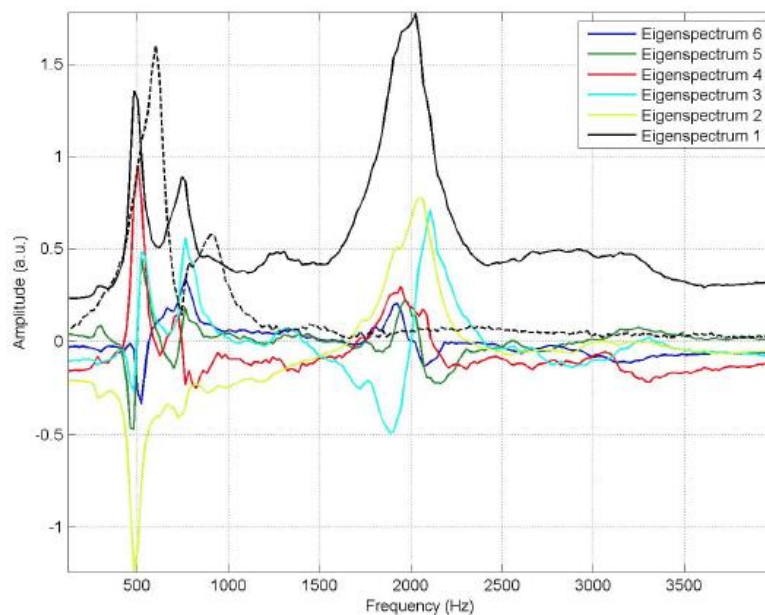


Рис.13. Усереднені спектри (eigenspectra).

При порівнянні двох вуликів виявлено, що деякі частотні ознаки є загальними для двох наборів eigenspectra. Чотири піки частот ідентифіковані близько 500Гц, 750Гц, 2000Гц і 3200Гц. Пара Eigenspectra (№3 і №5) мають зсув частоти в двох піках в одному і тому ж (№ 3) або в протилежних напрямках (№ 5), в той час як інша пара (№ 2 і 4) вказує амплітуди піків що регулюються незалежно. Деякі легкі відмінності також спостерігаються, ймовірно, через комбінацію природних резонансів вулика. На рисунку 13 пунктирною лінією наведено механічні резонанси вулика.

1.3. Виявлення статистично однорідних ділянок шуму бджолої сім'ї для діагностики стану колонії

Тобоев В.А [13], [14], вказував на розбіжність результатів акустичної діагностики у роботах інших дослідників, через використання методів аналізу, які не в повній мірі відображають зв'язки між змінами структури шуму і поведінки бджіл. Він вважав що застосування для виділення спектральних складових фільтрації акустичних шумів не дозволяє відокремлювати звуки, що генеруються бджолами, що готуються до ройового стану, від безлічі їх інших звуків. Так як важко відокремити звуки бджіл, які готуються до розмноження, від звуків бджіл, що знаходяться в різних поведінкових контекстах, то вплив останніх на знаходження статистично однорідних ділянок повинно бути зведений до мінімуму.

Приклад виділення статистично однорідних ділянок, які несуть в собі діагностичну інформацію про ройовий стан бджіл, показаний на рисунку 18 (ділянки виділені прямокутниками). Автор вважає що складені таким чином фрагменти дозволяють з високою надійністю визначати ступінь готовності до вильоту з вулика нової сім'ї.

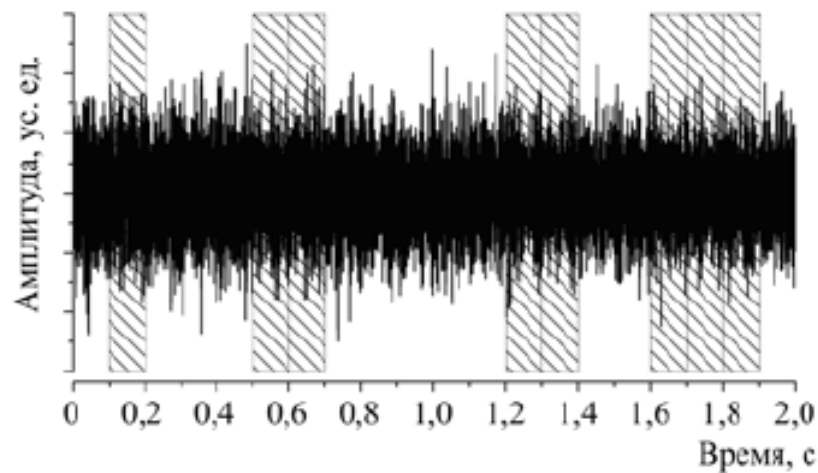


Рис.18. Виділення статистично однорідних ділянок акустичного шуму, що несуть в собі діагностичну інформацію (за 3 дні до виходу рою).

Дослідження функції загального середнього для акустичного шуму показує, що з наближенням вильоту рою підвищується ймовірність генерації бджолами статистично однорідних ділянок шуму. На це вказує збільшення загальної довжини таких ділянок і оптимальної ширини ділянок розбиття акустичного шуму на фрагменти. Наприклад, за 10 днів до виходу рою зазначена величина дорівнювала 48 мс, а вже за день до виходу рою збільшилася до 128 мс. Загальна довжина статистично однорідних ділянок збільшилася в середньому з 0.3 до 0.5 с в кожному зразку акустичного шуму. Автор вважає, що це пов'язано з синхронізованим рухом крил безлічі бджолиних особин, що складають основу майбутнього рою. Таку ж динаміку зміни шуму бджіл підтверджує аналіз трендів і функцій узагальненого середнього. До моменту виходу рою інтегровані послідовності виділених ділянок акустичного шуму стають детрендованими тобто розкид значень J_s щодо осі абсцис незначний. Цей факт зображений на рисунку 19(а).

Динаміка флуктуацій звукового шуму, представлена функціональною залежністю $Y_n(Y_0)$. Що до еталонного значення Y_0 , простежується збільшення степені кореляції з наближенням дня виходу рою (Y_n - відносні флуктуації за n днів до виходу риючи ($n = 30$)).

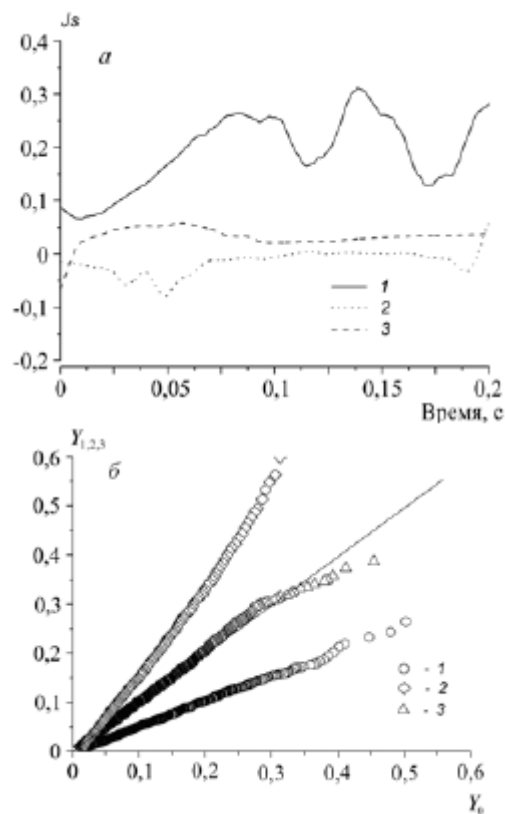


Рис.19. Динаміка зміни усереднених значень (трендів J_s) проінтегрованої послідовності виділених (а) для виділених ділянок акустичного шуму і взаємна кореляція. Послідовність ранжируваних амплітуд відносних флуктуацій (б). 1 - сім'я в активному (не ройовому) стані; 2 - за 10 днів до відльоту ройових бджіл; 3 - за день до відльоту.

Y_0 - відповідає послідовності ранжируваних амплітуд відносних флуктуацій (ПРАОФ) акустичного шуму в день виходу рою, числові характеристики якого отримані статистичною обробкою 30 сімей, які випустили рій. Надійність аналізу можна оцінити по відносній величині тангенса кута нахилу прямої по відношенню до ПРАОФ, віднесеного до еталонних значень ПРАОФ Y_0 (рис. 19, б). Так як для функції Y_n (Y_0) різняться ФОС для трендів і відносні флуктуації шуму мають також різні кількісні характеристики, їх можна використовувати для ідентифікації та кількісної диференціації ройового стану бджіл.

Початок періоду розмноження визначається по значному збільшенню кута нахилу відповідної прямої, значення тангенса якого стає більше одиниці (рис. 19, б). Динаміка зміни цього параметра така, що з наближенням дня виходу рою він прагне до одиниці. За п'ять днів до відділення сімей середнє значення тангенса кута нахилу прямих дорівнювало $1,12 \pm 0,4$, а за три дні $-1,04 \pm 0,02$. Для всіх восьми сімей, які відпустили рої, величина відсічення (значення ординати прямої в нульовій точці) в ці дні не перевищувала $0,02 \pm 0,01$. З порівняння залежності Y_n (Y_0) можна виділити ще два параметри: величину відносної похибки підгонки прямої (δ), яка в ідеалі близька до нуля, і

узагальнену функцію кореляції Пірсона (ОФКП). Ці чотири кількісні характеристики шуму утворюють кластер параметрів змінюються по відношенню до наближення виходу рою.

Для наочності інтерпретації і візуалізації ройового стану по знайденим кількісним характеристикам побудований еліптичний класифікатор, зображений на рисунку 20.

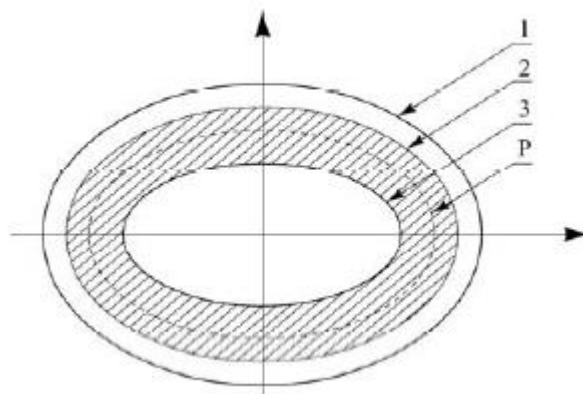


Рис.20. Еліптичний класифікатор визначення ройового стану бджіл по акустичному шуму. Еліпси побудовані з порівняння ФОС і ПРА відносних флуктуацій акустичного шуму в залежності від близькості моменту виходу рою: 1 - за 20 днів до виходу риючи; 2 - за 15 днів; 3 - день виходу рою. Заштрихована область відповідає акустичному шуму бджолою сім'ї, яка готується до розмноження.

1.4. Використання алгоритмів машинного навчання для моніторингу бджолиних сімей

У дослідницькій роботі А. Роблес-Герреро, Т. Соседо-Аная, Е. Гонсалес-Рамереза [15] було запропоновано методологію на основі MFCC та Machine Learning для автоматичного розпізнавання стану медоносної бджолиної колонії на основі аналізу звукозаписів. Система моніторингу розроблена на одноплатному комп'ютері Raspberry Pi 2В. Зразки звуку були отримані за допомогою всенаправлених мікрофонів що були розміщені всередині вуликів. Мікрофони були захищені металевою сіткою, щоб уникнути їх покриття воском та прополісом. Сигнал був прийнятий і перетворений в цифровий сигнал за допомогою мікроконтролера з dsp ядром та 12-бітним АЦП. Звукові дані збирали з двох колоній медоносних бджіл породи *Apis mellifera carnica*. Перша колонія мала королеву з великою кількістю бджіл. Друга колонія була без королеви зі зменшеною кількістю бджіл. Частота дискретизації вибірки для цього дослідження була встановлена 4 кГц, для зменшення частотного діапазону отриманих сигналів згідно теореми Найквіста. Кожен запис тривав 3 хв. Записи були робилися кожні 15 хв протягом 24 годин. Експеримент для збору даних проводили протягом 45 днів (з середини квітня до травня). На рисунку 14 показана спектрограма здорової колонії;

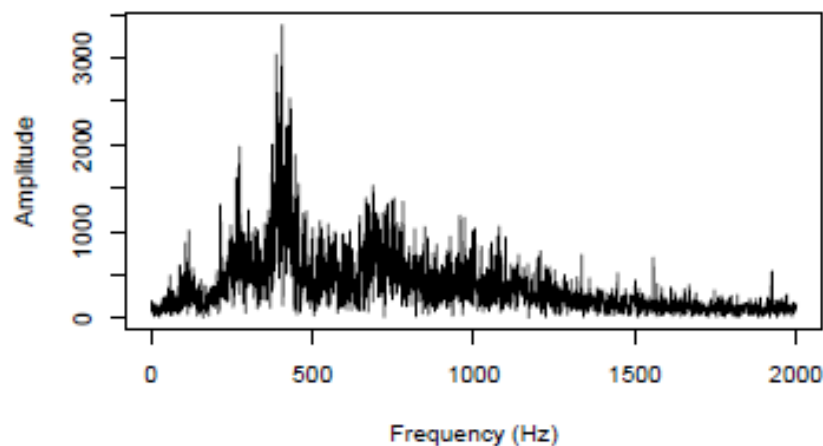


Рис.14. Частотна характеристика здорової колонії.

Більша частина інтенсивності розташована на смузі 400 Гц. На рисунку 15 колонія без королеви має інші частотні характеристики.

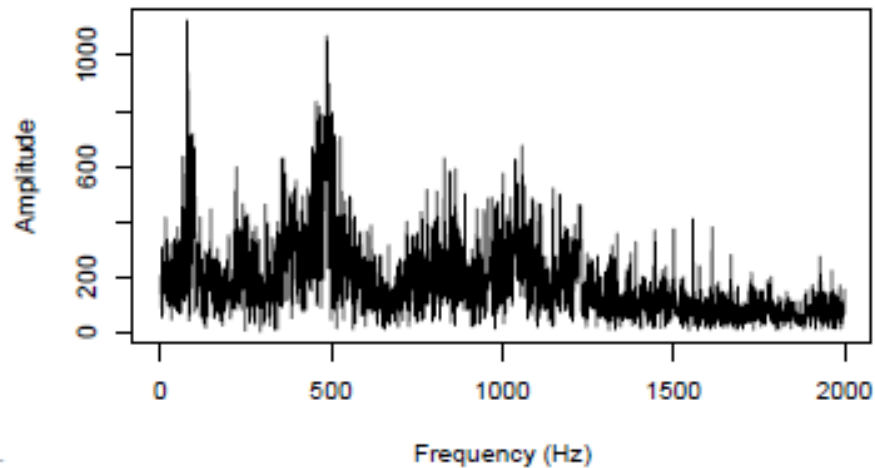


Рис.15. Частотна характеристика нездорової колонії

З рисунку 15 видно що для нездорової колонії звуковий спектр розподіляється за амплітудою в більшій кількості частотних смуг. Вибірка для MFCC складалася з 88 записів звуку, для здорової колонії з королевою і 98 записів для колонії без королеви. Ці дані зібрані з 24 годин запису.

Параметри розрахунку MFCC були наступними: розмір вікна 25 мс (у цей період часу сигнал вважається квазістаціонарним), і з 10 мс перекриття, значення попереднього нахилу було встановлено на 0,94. Ці значення зазвичай використовуються для розпізнавання мовлення. Два класи були обрані в цій роботі: здорова і нездорова колонія.

Перший проведений аналіз був розкладанням сингулярної величини (SVD). SVD обчислює набір власних значень і власних векторів матриці. Це звичайна методика аналізу багатовимірних даних, яка може виявити структури в наборі даних і використовується для проведення класифікації. Аналіз SVD набору даних показаний на рисунку 16.

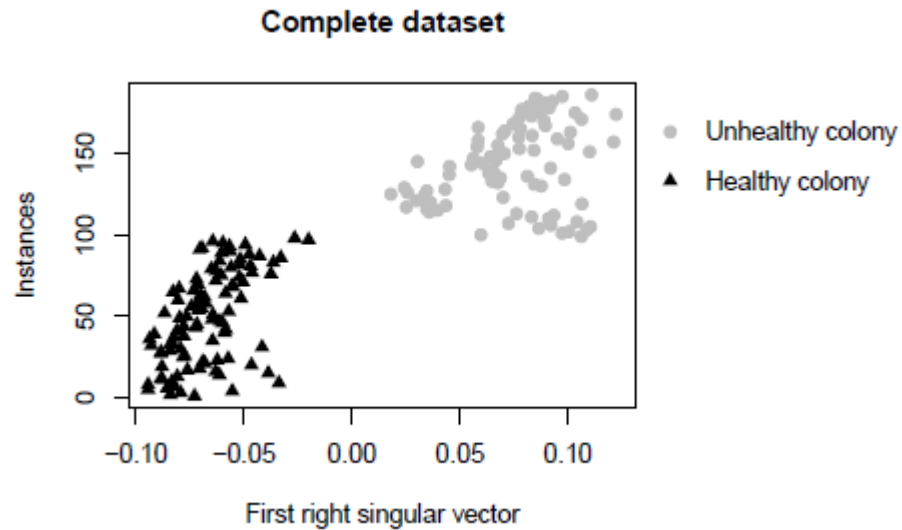


Рис.16. Класифікація за допомогою SVD

Чорний трикутник відповідає колонії з королевою (здорова), а сірий круг - колонія без королеви (нездорова). Зразки групуються, формуючи добре сформовані кластери, що свідчить про дві чітко помітні умови.

При візуальному обстеженні було визначено, що одного статистичного дескриптора достатньо для гарного прогнозування. Для оцінки характеристик моделі була отримана крива ROC, що показана на рисунку 17. Отримані операційні характеристики (ROC), є відомою моделлю вимірювання продуктивності для алгоритмів машинного навчання.

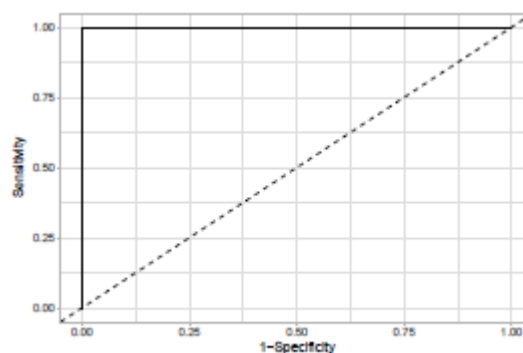


Рис.17. Крива ROC для даної системи класифікації.

Крива ROC показує кількість правильно класифікованих зразків у порівнянні з кількістю неправильно класифікованих негативних зразків.

Ідеальна класифікація відбивається кривою, яка лежить у верхньому лівому куті з унітарною областю під кривою ROC. Результатами даного дослідження є ефективна система класифікації (згідно такого показника як крива ROC), що за двома ознаками може зробити висновок здорова бджолина колонія чи ні.

Реалізацією системи акустичного моніторингу здоров'я бджолиних сімей у своїй дослідницькій роботі також займалися А. Кандар, І. Ахмад, Д. Хабібі і М. Леппард [16]. Дана система повинна розпізнавати бджолині сім'ї заражені кліщем Varroa (*Varroa destructor*). Запропонована система моніторингу складається з однопалатного комп'ютера Beagleboard та датчиків. Акустичні дані отримують за допомогою електретного мікрофону.

Звук бджолиних сімей, отриманий з мікрофона, обробляється beagleboard з використанням алгоритмів класифікації, щоб розрізнити здоровий або інфікований кліщем вулик. Радіопередавач малої потужності Zigbee Link використовуються для передачі даних у вигляді повідомлень про тривогу, у випадку виявлення інфекції, або про діагностичний стан системи.

Структурна схема системи, що використовується для аналізу акустичних зразків колонії медоносних бджіл зображена на рисунку 20.

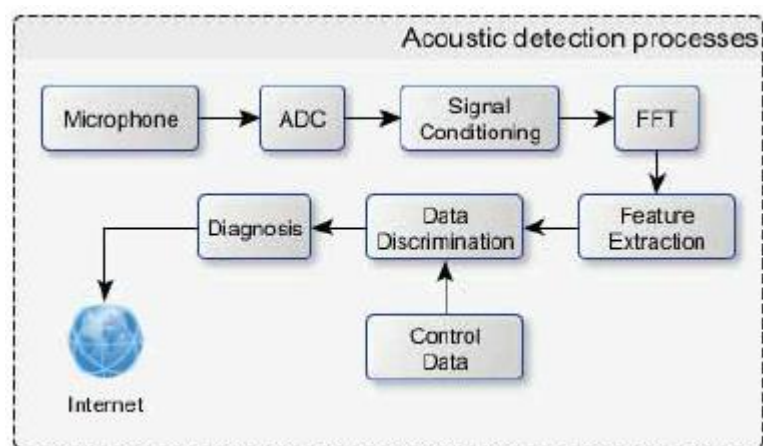


Рис.20. Структурна схема системи моніторингу бджолиних сімей

Навчальна вибірка складалась з звукових зразків здорової та інфікованої кліщем колонії. Кожен запис тривав 10 секунд. Акустичне порівняння між навчальною вибіркою і тестовою вибіркою дозволило дослідникам визначити,

що акустичний відбиток здорової колонії корелює з акустичними відбитками колонії зараженою шкідниками.

Методи акустичного аналізу які використовуються для визначення акустичних властивостей колонії вулика наступні:

- Peak Frequency (PF): PF можна визначити як смугу частот що має найбільшу потужність.
- Спектральний центроїд (SC): SC також відомий як середнє значення частоти або центр тяжіння спектру потужності.
- Пропускна здатність (B): смуга пропускання сигналу - це діапазон частот, присутніх в сигналі.
- Частота кореневих відхилень (RVF): функція RVF описує зближення потужності спектру між зразками що порівнюються.

Інструменти класифікації, що використовуються для виконання даного аналізу:

- Аналіз основних компонентів (PCA): PCA є інструментом аналізу даних, що використовується для прогнозування моделі. Зазвичай реалізується як форма

зменшення розмірності, що включає фіксацію власних значень і власних векторів коваріаційної матриці.

- Підтримуючі векторні машини (SVM): SVM є моделлю машинного навчання, побудована навколо алгоритмів, розроблених для аналізу даних і розпізнавання шаблонів.
- Лінійний дискримінантний аналіз (LDA): LDA використовується для пошуку лінійної комбінації набору ознак, що максимізує роздільність між класами.

Класифікація навчальної вибірки за допомогою PCA та RVF зображена на рисунку 21.

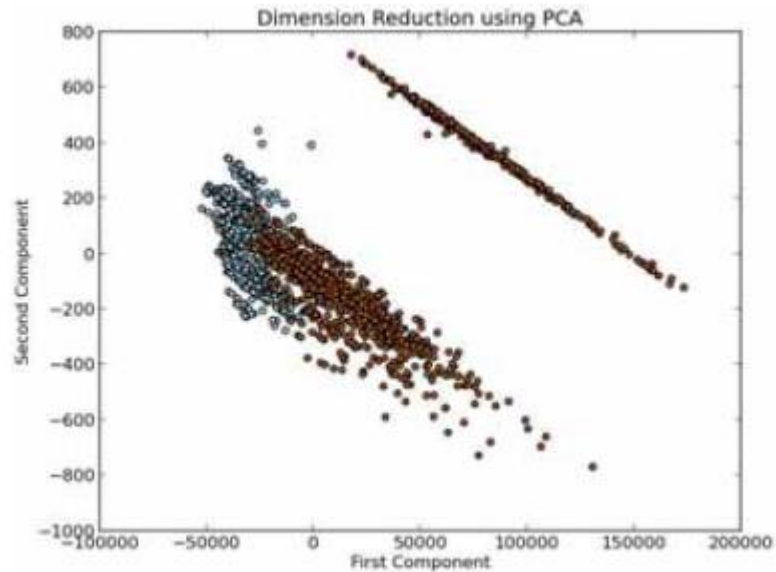


Рис.21. Аналіз PCA: здорові колонії (синій колір) – інфіковані (червоний); перший компонент (RVF) – другий компонент (B).

З рисунка 21 слідує що існує чіткий поділ (тобто мінімальне перекриття між кольоровими регіонами) між даними зі здорових та інфікованих кліщем Варроа колоній. Аналіз підтверджує придатність PCA для визначення аномалій в компонентах акустичного спектру (тобто розділення здорових і інфікованих ознак). Наступний етап передбачає визначення лінійної залежності (функції класифікатора) що найкраще описує найвищий ступінь роздільності між ними з використанням основних компонентів обраних PCA. Після знайдення відповідного відношення, набір тестових даних надходить у функцію класифікатора щоб зробити прогнози щодо стану бджолиної колонії. Класифікаційний аналіз навчальної вибірки з використанням лінійного дискримінантного аналізу (LDA) показаний на рисунку 22. Пунктирна лінія показує поріг класифікації на основі вибраних ознак (тобто B & RVF)

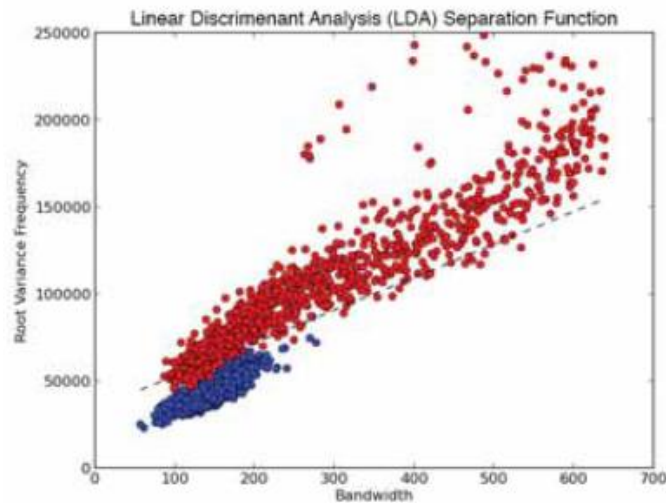


Рис. 22. Показана залежність дискримінантної функції що генерується алгоритмом LDA. Результати показують розмежування між здоровими (синій колір) і інфікованими (червоний) колоніями

Класифікація тестової вибірки за допомогою PCA та RVF зображена на рисунку 23.

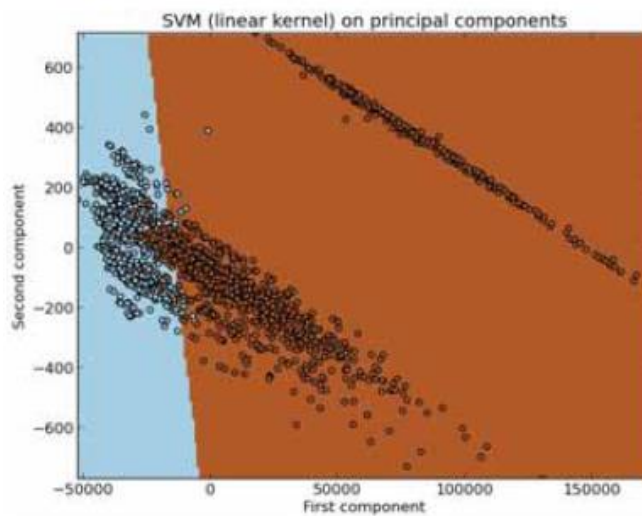


Рис. 23. Зображення здорових колоній (синій) проти інфікованих (червоний) використовуючи SVM - перший компонент (RVF) / другий компонент (B)

Для аналізу з використанням лінійного дискримінантного аналізу (LDA) було використано тестова вибірки що склалися з звукозаписів з двох

бджолиних сімей з невідомою мірою зараження. Класифікаційний аналіз тестової вибірки що складалась з першої сім'ї, з використанням лінійного дискримінантного аналізу (LDA) показаний на рисунку 24.

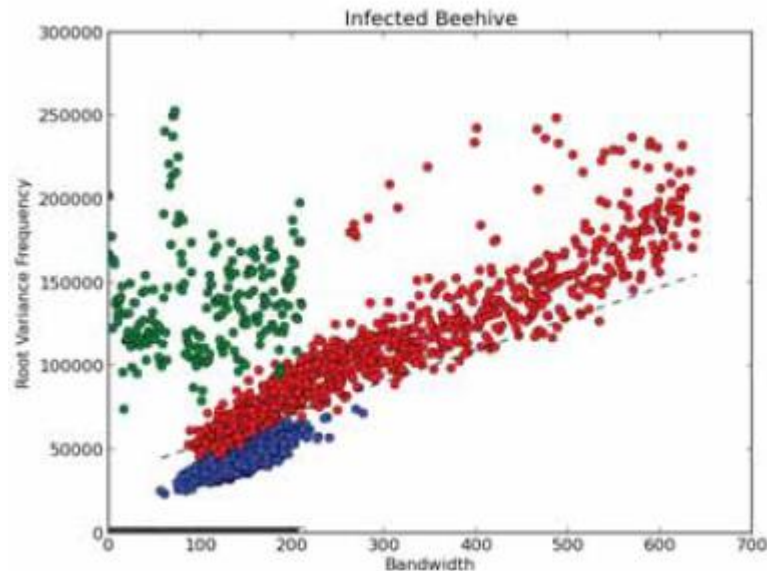


Рис. 24. Класифікація за допомогою LDA: здорові колонії – синій колір
інфіковані – червоний, тестова вибірка – зелений.

Синьо-червоні маркери використовуються для диференціації акустичних характеристик між здоровими та інфікованими колоніями для навчальної вибірки. Зелені маркери представляють набір даних від колонії з невідомим станом. При візуальному огляді є чітка вказівка на те, що тестова вибірка знаходиться в основному в межах зараженої зони. при детальному огляді даної колонії бджолярем було підтверджено що колонія була інфікована кліщем. Класифікація тестової вибірки що складалась з звукозаписів з другої сім'ї зображена на рисунку 25.

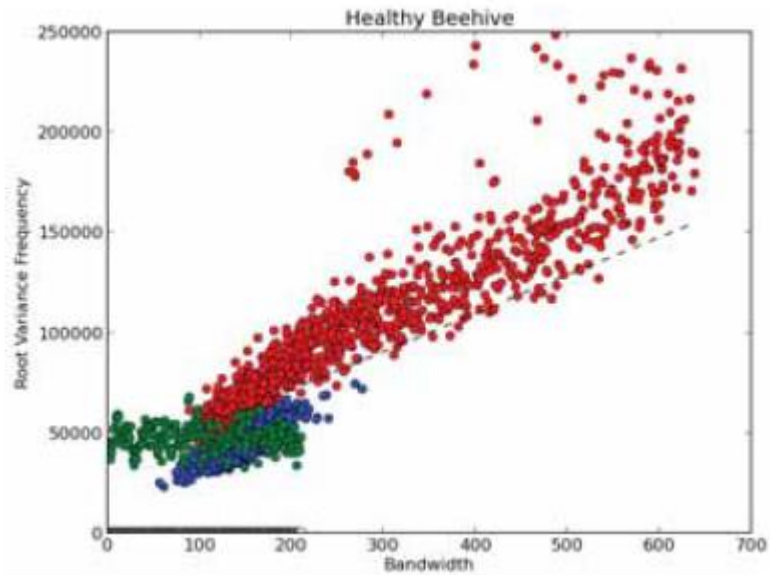


Рис. 25. Класифікація другої тестової вибірки за допомогою LDA: здорові колонії – синій колір інфіковані – червоний, тестова вибірка – зелений.

Розташування другої тестової вибірки на рисунку вказує на те, що друга колонія була здоровою, що було підтверджено при детальному огляді колонії бджолярем. Отже дослідниками була створена ефективна система класифікації що здатна відрізнити заражені кліщем Варроа колонії від здорових.

РОЗДІЛ 2. ОРГАНІЗАЦІЯ ЕКСПЕРИМЕНТАЛЬНИХ ДОСЛІДЖЕНЬ ДЛЯ ЗБОРУ ДАНИХ. ПОПЕРЕДНЯ ОБРОБКА ДАНИХ.

2.1. Загальний опис та мета проведених досліджень

В даній дослідницькій роботі при розробці експериментальних досліджень ставилася наступна класифікаційна задача: знайти частотні складові звукового спектру бджолої сім'ї, які дозволять надійно відрізняти такі стани бджолої сім'ї як: стан колонії під час медозбору і його відсутності та стан колонії під час втрати королеви. Також класифікаційна задача включає в себе розробку правил прийняття відповідних рішень на основі отриманих інформаційних частотних складових. Отримані результати дозволили б у майбутньому створити програмно-апаратну систему класифікації зазначених станів бджолої сімей породи *Apis mellifera sossimai*.

Дана дослідницька робота містила у собі чотири експериментальних дослідження. Два з них спрямовані на створення системи класифікації що дозволить відрізняти стан колонії під час медозбору від стану під час відсутності медозбору за допомогою частотних складових звукового спектру. Інші два експериментальні дослідження спрямовані на створення системи класифікації що дозволить відрізняти стан колонії під час втрати королеви від нормального стану колонії, тобто стану під час відсутності медозбору але присутності в колонії королеви.

2.2. Організація першого експериментального дослідження

Перший експеримент проведений в даному дослідженні був націлений на збір даних, у вигляді аудіо записів, для перевірки взаємозв'язків частотних складових звукового спектру вулика з станом бджолої сім'ї під час втрати королеви. В експерименті брало участь шість вуликів-нуклеусів, приблизно однакового розвитку, породи *Apis mellifera sossimai*. В перших трьох нуклеусах були матки однакового походження і віку. В інших трьох нуклеусах

матки були забрані в один і той же день. Кожен нуклеус мав свою кодову назву. Експеримент проводився з 29.06.2017 по 9.07.2017 включно. Тобто в другій трійці нуклеусів матки були відсутні впродовж 11 днів. За цей час бджоли створили маточники, з яких повинні вийти нові матки. Впродовж експерименту маточники не видалялися. Цей стан сім'ї не можна вважати ройовим, так як під час роїння стара матка сама відкладає яйця в мисочки, зі створених маточників формуються нові матки, а стара матка йде у рій. Бджолина сім'я яка йде в рій має особливий стан, діяльність бджіл всередині колонії порушується, медозбірна активність падає, так як льотна бджола готується йти в рій. Це звісно позначається на частотних складових звукового спектру [9]. В результаті виходу рою в сім'ї залишається тільки одна молода матка, яка вийшла останньою. Інші молоді матки також ідуть у рій. В даній ситуації роїння не було так як стара матка втрачена, тому молода матка яка вийде першою знищить інших молодих маток. Експеримент був завершений до моменту виходу маток так як нас цікавить саме стан втрати королеви, та система класифікації що дозволить виявити втрату якнайшвидше. Слід зазначити що під час експерименту в трьох нуклеусів з бджолиними королевами було зафіксовано стабільний поступовий розвиток, без будь-яких порушень, що могли значно вплинути на експеримент. Під час експерименту були зафіксовані такі параметри як: температура повітря, погодні умови, точний час початку і закінчення експерименту. Дані параметри вказані в таблиці 1.

Табл. 1. Умови проведення першого експерименту.

Дата	Температура повітря	Час початку експерименту	Час кінця експерименту	Погодні умови
29 червня	26°C	09:01	09:29	Сонячно
30 червня	26°C	09:01	09:28	Хмарно
01 липня	28°C	09:00	09:24	Солнечно

02 липня	23°C	08:58	09:28	Хмарно, перед дощем
03 липня	25°C	09:05	09:30	Помірна хмарність
04 липня	21°C	12:10	12:40	Хмарно, після дощу
05 липня	20°C	09:00	09:28	Сонячно
06 липня	24°C	09:10	09:40	Сонячно
07 липня	17°C	08:50	09:20	Сонячно
08 липня	19°C	09:00	09:30	Сонячно
09.липня	14°C	09:03	09:30	Хмарно, після дощу

Для всіх записів частота дискретизації становить 44100Гц, формат WAV. Записи були зроблені на MEMC-мікрофон, що вбудований в смартфон Meizu M2 mini. Розміщення мікрофона всередині нуклеуса було рівновіддаленим відносно стін нуклеуса. Верхня кришка та поточини вулика були зняті в ході експерименту так як смартфон був зафіксований вручну. Розміщення смартфона в колонії спричиняло незручності для бджіл на якийсь час. Це звісно позначалося б на спектральних характеристиках отриманих записів. Тому в усіх записах були вручну видалені фрагменти коли бджоли оточували мікрофон, спричиняючи небажані шуми. Смартфон було розміщено між рамками з печатним розплідом, напроти льотка вулика. Таке розміщення було вибрано тому що між розплідом знаходиться велике скупчення як бджіл годувальниць, так і льотних бджіл. Саме бджоли годувальниці мають створювати інформативні частотні складові для розпізнання стану втрати королеви [7]. Також слід зазначити що саме на рамках з розплідом створювались маточники. В результаті експерименту було отримано 33 записи для ситуації втратою королеви і 33 записи нормального стану колонії. Кожен запис тривав близько півтори хвилини.

2.3. Попередня обробка даних з першого експериментального дослідження

Попередня обробка даних отриманих в результаті першого експерименту була виконана за допомогою метода Уелча [17]. До застосування методу Уелча записи були пронормовані. У методі Уелча сигнал який аналізується, представлений вектором відділків, розбивається на сегменти що перекриваються. Кожен сегмент помножується на вагову функцію. Використання вагової функції дозволяє послабити розтікання спектра і зменшити зміщення отриманої оцінки спектра густини потужності ціною незначного погіршення роздільної здатності. Потім для кожного сегмента обчислюється періодограма, потім отримані періодограми усереднюються. В даному дослідженні для обчислення періодограм використовувалася вагова функція Кайзера [18]. Метод Уелча дозволяє отримати графічне відображення розподілу оцінок густини потужності сигналу відносно частоти, тобто отримати усереднену періодограму. З кожного вулика було отримано 11 записів, і до кожного з них був застосований метод Уелча. На рисунках 26-31 представлено усереднені періодограми до кожного з записів (показані різним кольором) для кожного вулика, та усереднена по 11 отриманим періодограмам періодограма.

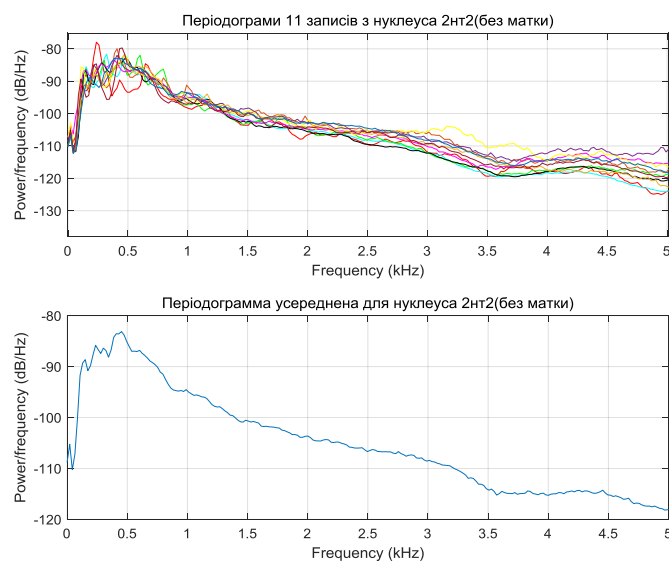


Рис. 26. Періодограми для нуклеуса 2нт2, королева відсутня.

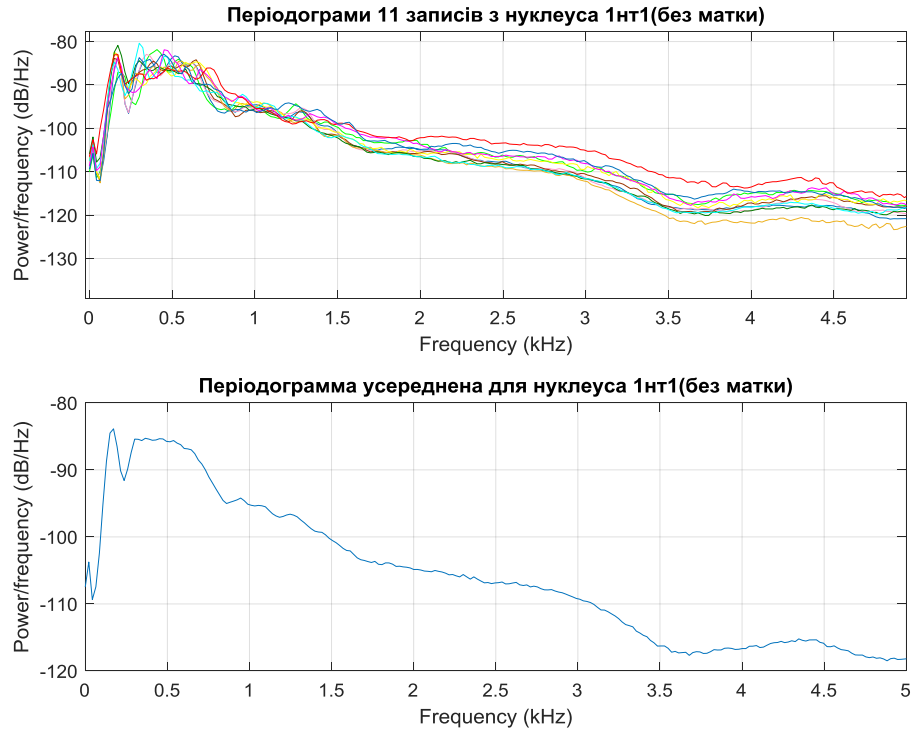


Рис. 27. Періодограми для нуклеуса 1нт1, королева відсутня.

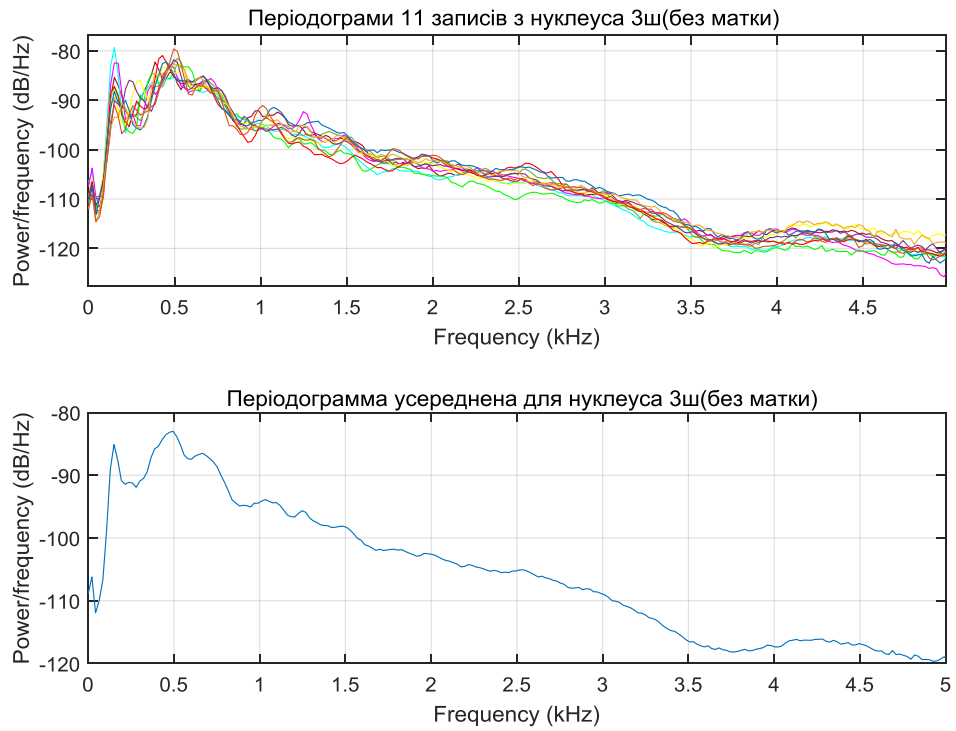


Рис. 28. Періодограми для нуклеуса 3ш, королева відсутня.

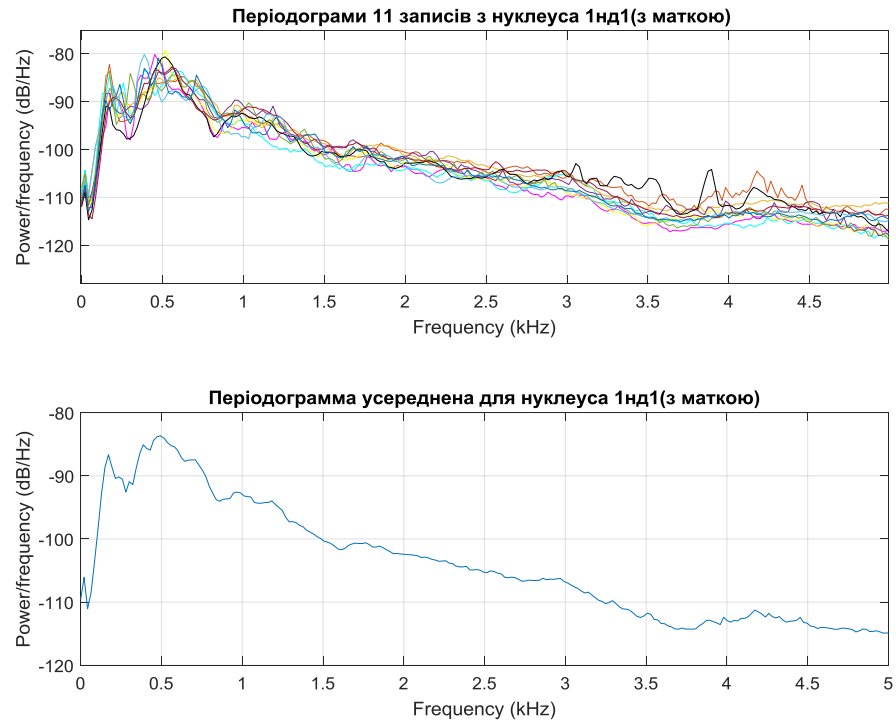


Рис. 29. Періодограми для вулика 1нд1, королева присутня.

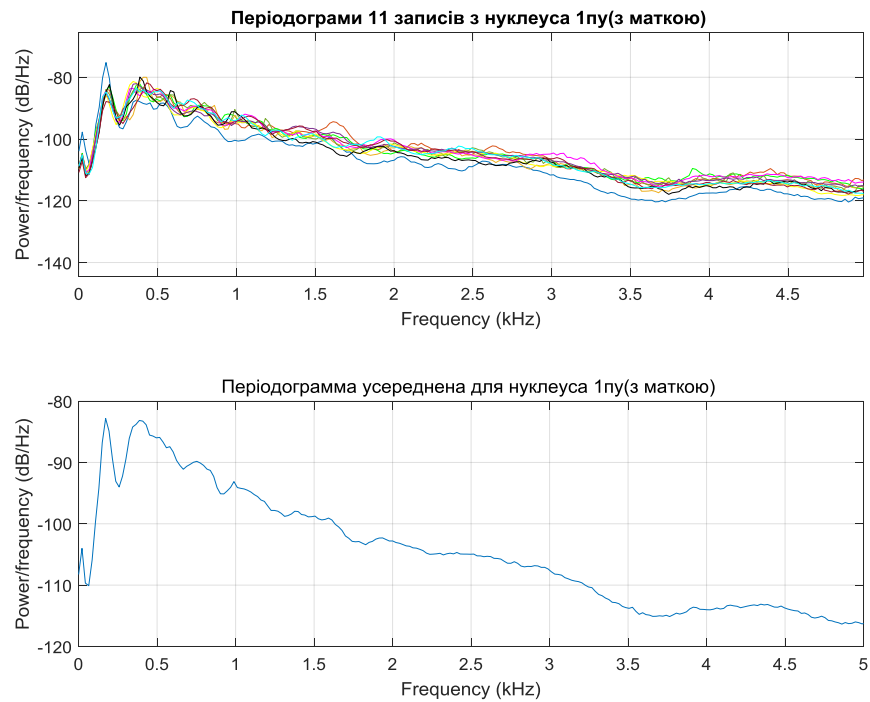


Рис. 30. Періодограми для вулика 1пу, королева присутня.

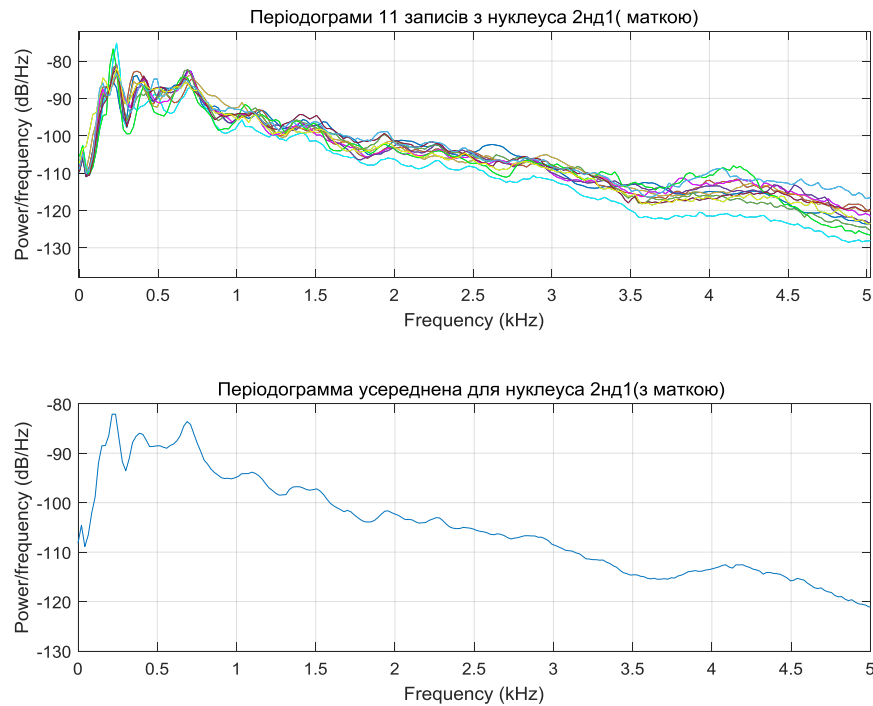


Рис. 31. Періодограми для вулика 2нд1, королева присутня.

Як видно з вищевказаних усереднених періодограм, максимуми оцінки густини потужності знаходяться на відносно різних частотах, тобто зміщуються, незалежно від класифікаційного стану колоній. Ще більші розбіжності між максимумами оцінки густини потужності відслідковуються на зображенні всіх 11 періодограм для кожного з вуликів. Можна зробити висновок що бджолина колонія це динамічна система, спектральні характеристики якої змінюються навіть за відносно однакових умов проведення експерименту. Також слід вказати що інформативний частотний діапазон можна скоротити до 1.5кГц.

Отримані для кожного з нуклеусів усереднені періодограми були знову усереднені, щоб отримати дві періодограми для різних класифікаційних станів. Отримані періодограми зображені на рисунку 32.

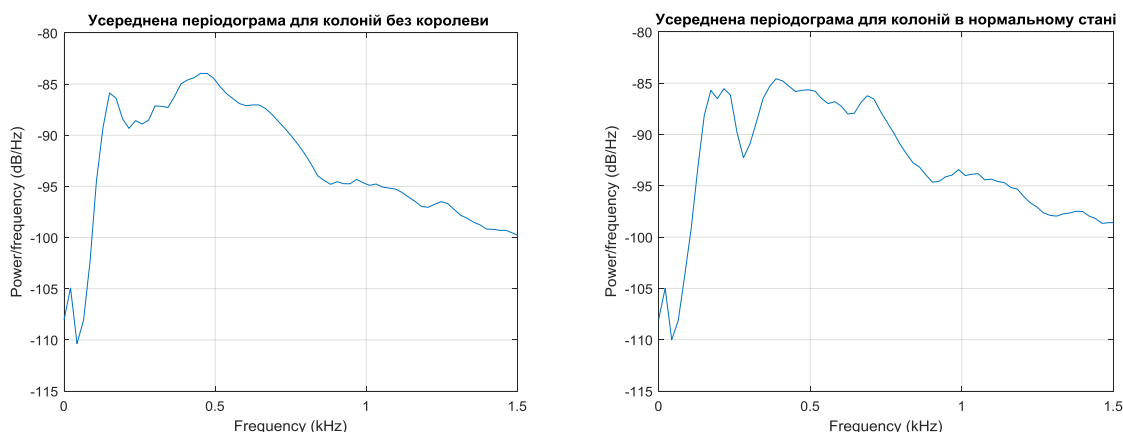


Рис. 32 .Усереднені періодограми для двох класифікаційних станів: присутності королеви у колонії та її відсутності.

Розглянувши рисунок 32 можна помітно що на періодограмі для нормального стану бджолої сім'ї максимумами оцінки густини потужності знаходяться на смугах частот 170-210Гц, 400Гц, та 700Гц в той час як для бджолиних сімей без королеви перший максимум оцінки густини потужності зсувається на частоту 150Гц, другий максимум зсувається на частоту 450Гц, а третій максимум зникає. Отримані спектральні характеристики суперечать результатам досліджень Е.К. Еськова [7] та С. Фераррі [9]. Отже дані отримані в результаті даного експерименту треба спробувати обробити іншими методами.

2.4. Організація другого експериментального дослідження

Другий експеримент був націлений на збір даних, у вигляді аудіо записів, для перевірки взаємозв'язків частотних складових звукового спектру вулика і активності бджіл під час медозбору. В експерименті брало участь три вулики, приблизно однакового розвитку, з племінними матками однакового походження породи *Apis mellifera sossimai*. Кожен вулик мав свій номер. Експеримент був поділений на два етапи. Перший етап проводився з 22.07.2017 по 28.07.2017 включно. У ньому були зроблені записи під час медозбору. Під час експерименту були зафіксовані такі параметри як:

температура повітря, погодні умови, точний час початку і закінчення експерименту, а також приріст маси вулика, що є головним показником медозбору на першому етапі. Дані параметри вказані в таблиці 2.

Табл. 2. Умови проведення другого експерименту, перший етап.

Дата	Температура повітря	Час початку експерименту	Час кінця експерименту	Погодні умови	Взяток за добу
22 липня	28°C	08:57	09:10	Сонячна	1.6 кг
23 липня	26°C	09:12	09:25	Хмарна	1 кг
24 липня	22°C	09:15	09:25	Сонячна	1.3 кг
25 липня	24°C	09:02	09:20	Сонячна	1.7 кг
26 липня	26°C	09:05	09:20	Сонячна	800 г
27 липня	27°C	08:55	09:15	Сонячна	700 г
28 липня	26°C	09:00	09:16	Хмарна	700 г

Другий етап проходив з 19.08.2017 по 27.08.2017 включно, в ньому були зроблені записи під час повного припинення медозбору. Записи були здійснені приблизно в один і той же час протягом всього експерименту. Для другого етапу також були зафіксовані вищевказані параметри, вони вказані в таблиці 3.

Табл. 3. Умови проведення другого експерименту, другий етап.

Дата	Температура повітря	Час початку експерименту	Час кінця експерименту	Погодні умови
19 серпня	23°C	08:46	09:00	Сонячна
20 серпня	21°C	08:48	09:05	Сонячна
21 серпня	16°C	08:45	09:00	Сонячна
22 серпня	23°C	08:50	09:05	Хмарна
23 серпня	16°C	08:50	09:05	Хмарна
24 серпня	16°C	09:00	09:12	Сонячна
25 серпня	18°C	08:54	09:10	Сонячна

вручну видалені фрагменти коли бджоли оточували мікрофон, щоб виключити вплив шумів які вони спричиняють. В результаті експерименту було отримано 27 записів для ситуації з повною відсутністю медозбору і 21 запис для ситуації медозбору. Кожен запис тривав близько півтори хвилини.

2.5. Попередня обробка даних з другого експериментального дослідження

Попередня обробка даних отриманих в результаті другого експерименту була виконана за допомогою метода Уелча. До застосування методу Уелча записи були пронормовані. В даному дослідженні для обчислення періодограм використовувалася вагова функція Кайзера [18]. Метод Уелча дозволяє отримати графічне відображення розподілу оцінок густини потужності сигналу відносно частоти, тобто отримати усереднену періодограму. З кожного вулика було отримано 9 записів, під час повного припинення взятку і 7 записів під час взятку. До кожного з них був застосований метод Уелча. На рисунках 34-39 представлено усереднені періодограми до кожного з записів (показані різним кольором) для кожного вулика, та усереднена по отриманим періодограмам періодограма.

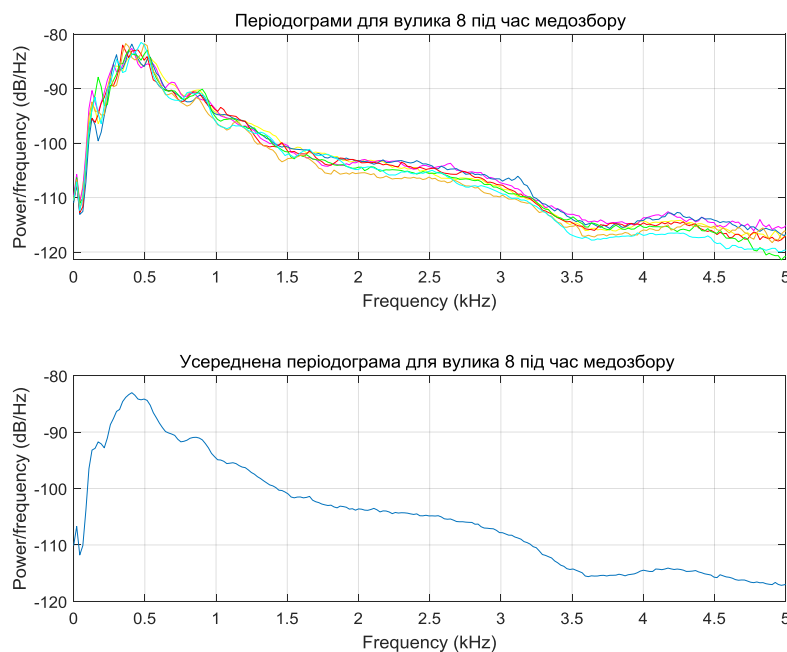


Рис. 34. Періодограми для вулика 8 під час медозбору.

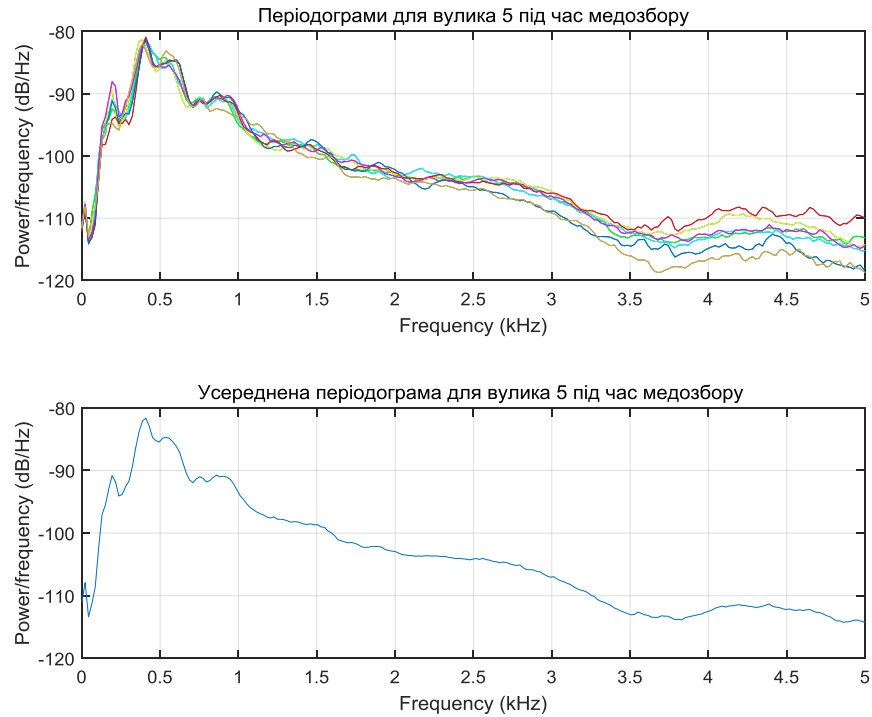


Рис. 35. Періодограми для вулика 5 під час медозбору.

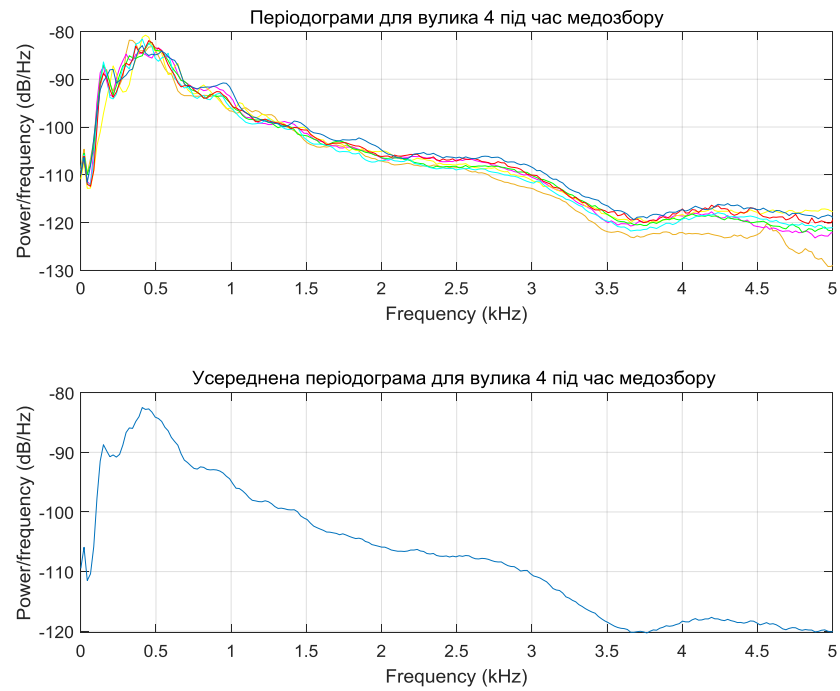


Рис. 36. Періодограми для вулика 4 під час медозбору.

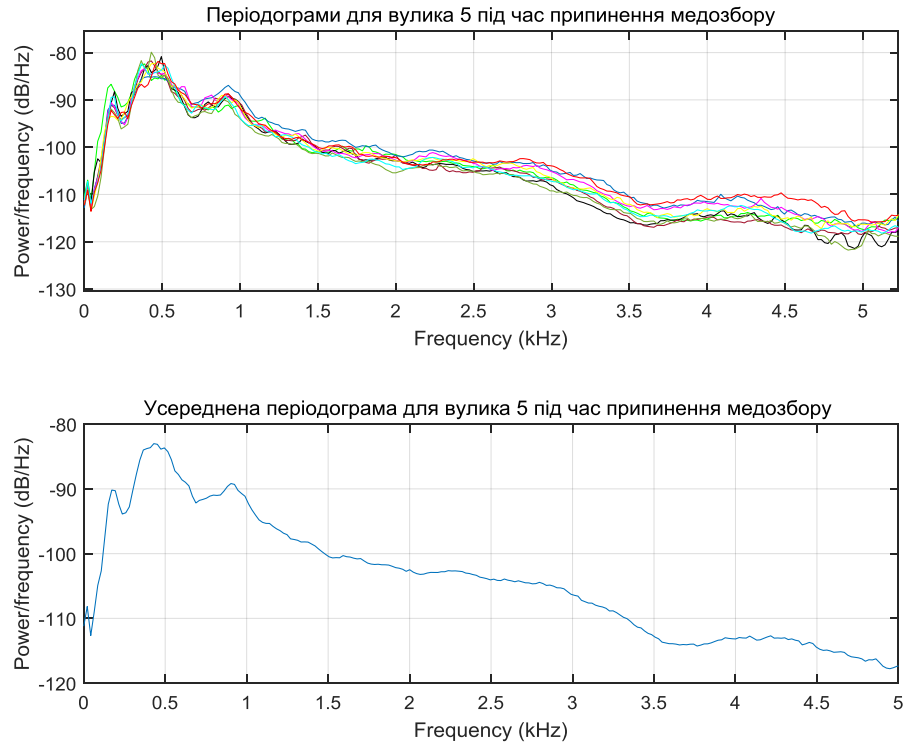


Рис. 37. Періодограми для вулика 5 під час припинення медозбору.

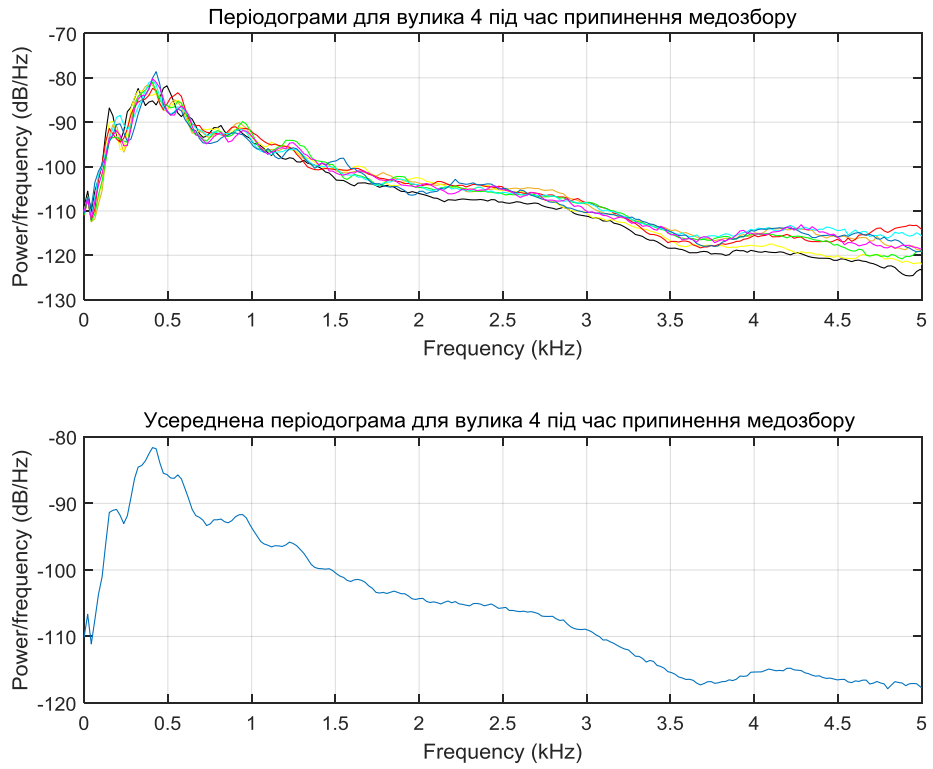


Рис. 38. Періодограми для вулика 4 під час припинення медозбору.

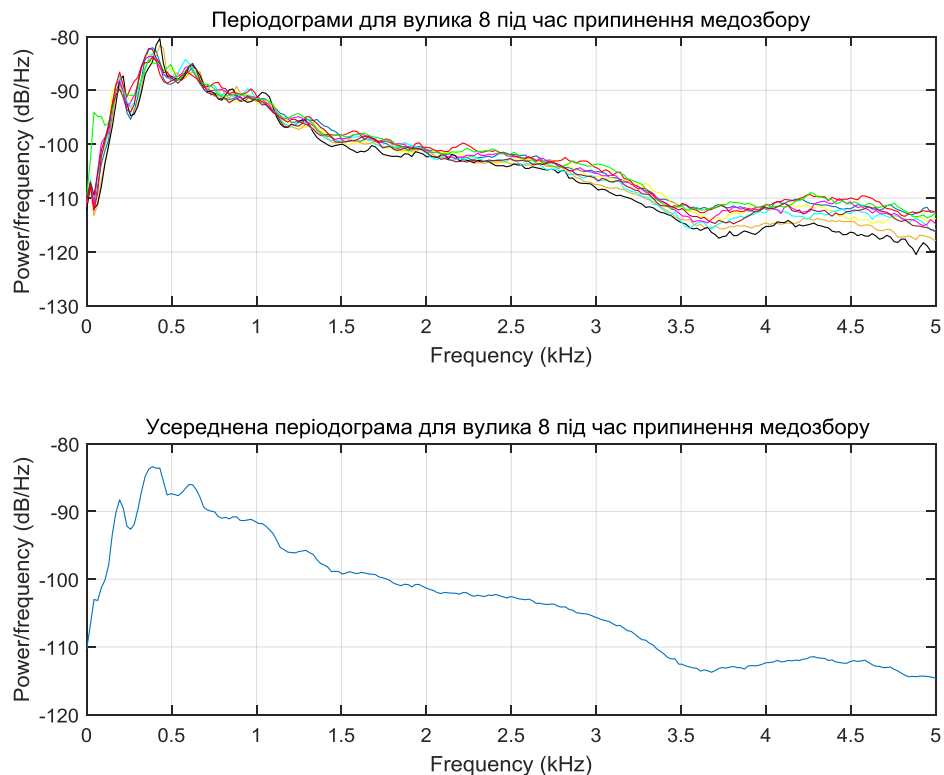


Рис. 39. Періодограми для вулика 8 під час припинення медозбору.

Як видно з вищевказаних усереднених періодограм, максимуми оцінки густини потужності знаходяться на відносно однакових частотах, незалежно від класифікаційного стану колоній. Максимуми оцінки густини потужності що відслідковуються на зображеннях всіх 11 періодограм, для кожного з вуликів здебільшого повторюються. Також слід вказати що інформативний частотний діапазон можна скоротити до 1.5кГц.

Отримані для кожного з вуликів усереднені періодограми були знову усереднені, щоб отримати дві періодограми для різних класифікаційних станів.

Отримані періодограми зображені на рисунку 40.

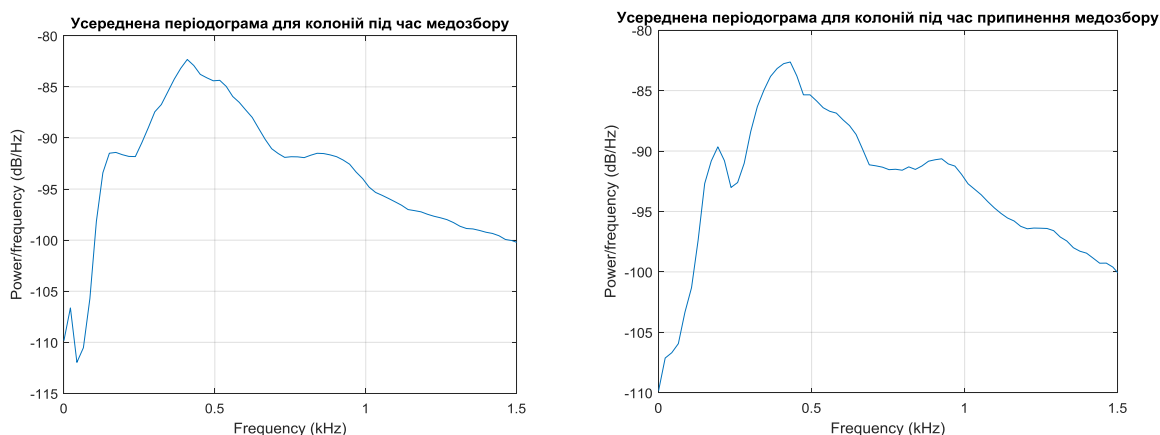


Рис. 40. Усереднені періодограми для двох класифікаційних станів: медозбору та його повного припинення.

Розглянувши рисунок 40 помітно що періодограми для даних двох класифікаційних станів відрізняються тільки тим, що на усередненій періодограмі, отриманій під час повного припинення медозбору, більш чітко виражений максимум оцінки густини потужності на частоті 200Гц. Слід зазначити що на обох періодограмах для даних класифікаційних станів присутній максимум оцінки густини потужності на частоті 400Гц. Отримані спектральні характеристики суперечать результатам досліджень Е.К. Еськова [7] та С. Фераррі [9]. Отже дані отримані в результаті даного експерименту треба спробувати обробити за допомогою інших методів.

2.6. Організація третього експериментального дослідження

Третій експеримент являвся модернізацією першого експерименту. Метою експерименту був збір даних, у вигляді аудіо записів, для перевірки взаємозв'язків частотних складових звукового спектру вулика з станом бджолої сім'ї під час втрати королеви. В експерименті брало участь дванадцять бджолиних сімей породи *Apis mellifera scutellata*. Бджолині сім'ї були приблизно однакового розвитку. Кожен вулик мав свій номер. Експеримент проводився з 27.04.2018 по 4.05.2018. З кожного вулика була вилучена королева, після чого були зроблені записи шуму кожної сім'ї, що

тривали 6 годин, із 13:30 до 19:30. Для подальшого аналізу використовувалися 12 відрізків 6-годинного запису, кожний тривалістю 6-7 хвилин, взяті через кожні півгодини. Тобто в порівнянні з першим експериментом вибірка була збільшена. Також слід зазначити що дана вибірка сформована таким чином щоб майбутня система класифікації могла виявити втрату королеви на самому початку цього стану, тобто до того моменту коли бджоли почнуть формувати маточники. Е.Ф. Вудс [4] у своїй дослідницькій роботі вказував що бджоли розпізнають втрату королеви через 3 години після вилучення, тому шести годин повинно бути достатньо щоб сформувати вибірку, за якою можливо розробити правило прийняття рішення, що забезпечить якісну класифікацію.

Під час експерименту були зафіксовані такі параметри як температура повітря та погодні умови. Дані параметри вказані в таблиці 4.

Табл.4. Умови проведення третього експерименту

Дата	Температура повітря	Номери вуликів, з яких були отримані дані.	Погодні умови
27 квітня	23°C	№2;	Сонячна
28 квітня	20°C	№1; №13;	Хмарна
30 квітня	24°C	№3; №4;	Сонячна
2 травня	23°C	№8;	Сонячна
3 травня	21°C	№9; №11;	Хмарна
4 травня	25°C	№4; №10;	Сонячна
7 травня	23°C	№6; №7;	Сонячна

Записи шуму зазначеної породи бджіл виконано із частотою дискретизації 44100 Гц за допомогою двох вимірювальних ненаправлених конденсаторних мікрофонів Superlux ECM-999 та зовнішньої двоканальної 24-бітної звукової карти Presonus Audiobox 96. Основні технічні характеристики мікрофону Superlux ECM-999 [20] вказані у таблиці 5.

Табл. 5. Технічні характеристики мікрофону Superlux ECM-999

Робочий частотний діапазон	20-20кГц
Вихідний імпеданс	200 Ом
Чутливість	14мВ/Па
Відношення сигнал-шум	70дБ
Динамічний діапазон	106дБ
Максимально допустимий рівень звукового тиску	132дБ

Амплітудно частотна характеристика мікрофона [20] наведена на рисунку 41.

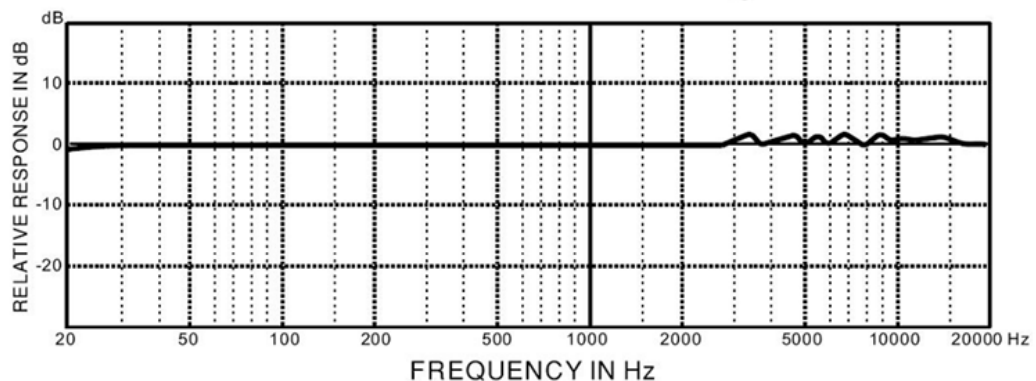


Рис. 41. АХЧ мікрофона Superlux ECM-999

Мікрофони були під'єднані до звукової карти за допомогою XLR кабелів. Така комутація забезпечує передачу корисного сигналу на значні відстані без спотворень. Слід зазначити що під час експерименту використовувались XLR кабелі довжиною більше 10 метрів. Схема системи зчитування даних зображена на рисунку 42.

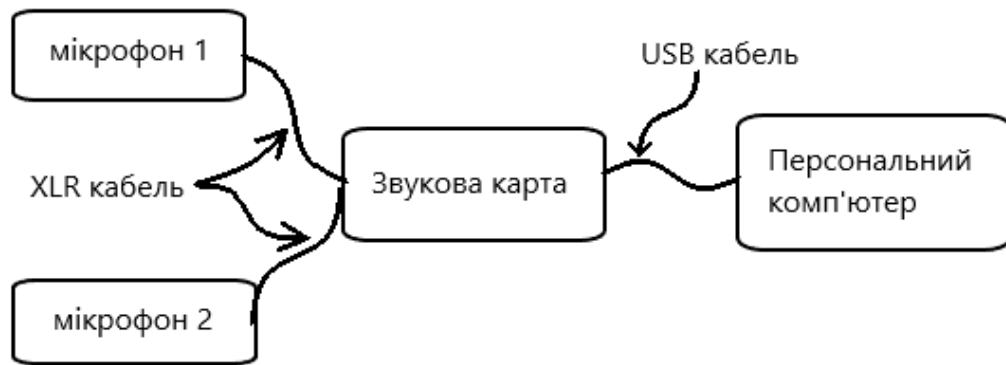


Рис. 42. Структурна схема системи запису даних.

Як видно з рисунку 42 така система забезпечувала запис аудіо даних з двох вуликів одночасно. Розміщення мікрофонів у вуликах було подібним до розміщення мікрофона у другому експерименті, рисунок 33. Мікрофон розміщувався у каналі між п'ятою і шостою рамкою від початку гнізда. П'ята і шоста рамки містили у собі велику кількість печатного розплоду. Причини з яких в даному дослідженні мікрофон було розміщено саме між рамками з розплодом на початку гнізда обґрунтовано у описанні першого експерименту. Розташування мікрофона відносно рамок та засіб фіксації мікрофона зображені на рисунку 43.



Рис. 43. Розміщення мікрофону відносно рамок і поточин

Таке розташування мікрофона між рамками дає змогу робити записи з закритими поточинами та кришкою, а значить значно зменшує втручання під час експерименту в екосистему вулика. В усіх записах були вручну видалені фрагменти коли бджоли оточували мікрофон, щоб виключити вплив шумів які вони спричиняють. В результаті експерименту було отримано 144 записи для ситуації з втратою бджолиної королеви.

2.7. Попередня обробка даних з третього експериментального дослідження

Для аналізу отриманих записів сигналів в діапазоні частот від 125 до 1250 Гц використано гребінку із одинадцяти 1/3-октавних нерекурсивних смугових фільтрів Чебишова другого роду [21]. Значення центральних частот f_0 та смуг пропускання Δf фільтрів наведено в таблиці 6. Порядок фільтрів становив 5000, що дозволило реалізувати амплітудно-частотну характеристику (АЧХ) фільтрів практично прямокутної форми (рисунок 44) із низьким рівнем бокових пелюстків (-110 дБ).

Табл. 6. Параметри гребінки фільтрів

f_0 , Гц	Δf , Гц
125	112-140
160	140-180
200	180-224
250	224-280
315	280-355
400	355-450
500	450-560
630	560-710
800	710-900
1000	900-1120
1250	1120-1400

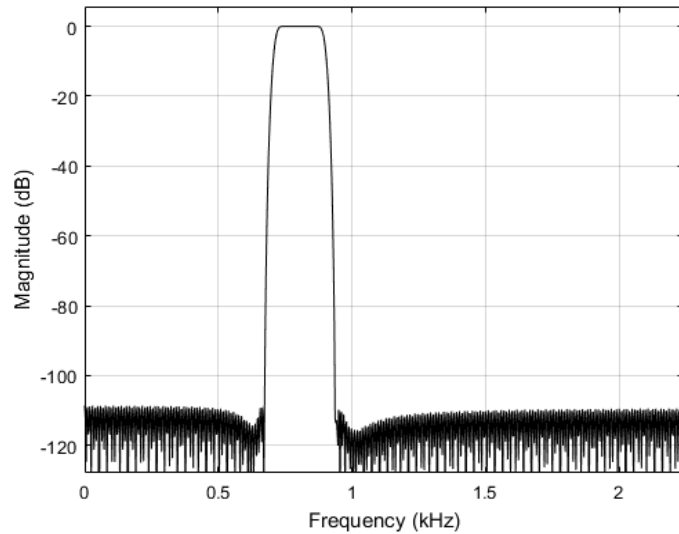


Рис. 44. АХЧ фільтра для $f_0=800\text{Гц}$

Шляхом оцінювання дисперсії D_i сигналів із виходів кожного i -го фільтра, було отримано значення G_i , $i = \overline{1,11}$ оцінок спектру потужності сигналів, що аналізуються:

$$G_i = \frac{D_i}{\Delta f_i} . \quad (1)$$

Для знаходження «первинних» класифікаційних ознак використано значення G_i , $i = \overline{1,11}$ нормовані за по дисперсією D процесу, що аналізується:

$$P_i = \frac{G_i}{D} . \quad (2)$$

2.8. Організація четвертого експериментального дослідження

Четвертий експеримент являвся модернізацією другого експерименту. Метою експерименту був збір даних, у вигляді аудіо записів, після обробки яких можливо буде знайти частотні складові звукового спектру бджолоїної сім'ї, які дозволять відрізнати два стани бджолоїної сім'ї – під час медозбору

та його відсутності. Записи шуму сімей породи *Apis mellifera sossimai* виконано із частотою дискретизації 44100 Гц за допомогою двох вимірювальних ненаправлених конденсаторних мікрофонів Superlux ECM-999 та зовнішньої двоканальної 24-бітної звукової карти Presonus AudioBox 96, що були використані у попередньому експерименті. Структурна схема системи запису даних показана на рисунку 42. Розміщення мікрофонів у вулику було виконано так само, як в третьому експерименті (рисунок 43), та забезпечувало значно зменшує втручання під час експерименту в екосистему вулика, в порівнянні з другим експериментом. Мікрофон розміщувався у каналі між п'ятою і шостою рамкою від початку гнізда. Як і в другому експерименті даний експеримент був поділений на два етапи. Перший етап (медозбір) проводився з 12.07.2018 по 23.07.2018. Другий етап (припинення медозбору) проводився з 19.08.2018 по 22.08.2018. Під час медозбору записи було отримано з 7 бджолиних сімей, а під час повного припинення медозбору записи було отримано з 8 бджолиних сімей. Бджолині сім'ї, які були задіяні в експерименті, були приблизно однакового розвитку, з бджолиними матками однієї породи. Запис шуму кожної сім'ї тривав 6 годин, із 13:30 до 19:30. Для подальшого аналізу використовувалися 12 відрізків 6-годинного запису, кожний тривалістю 6-7 хвилин, взяті через кожні півгодини.

Погодні умови під час проведення експерименту були приблизно однаковими для обох ситуацій, що розглядалися. Приріст маси вулика на контрольних вагах, що є головним показником медозбору, коливався від трьох до п'яти кілограм за день. Слід зазначити що в даному експерименті в порівнянні з другим експериментом записи проводились в степу, тобто в середовищі яке передбачає набагато вищу медозбірну активність бджіл. Показником який на це вказував був приріст маси вулика на контрольних вагах, що був в два-три рази більший ніж у другому експерименті. Також у порівнянні з другим експериментом була збільшена вибірка. В результаті

експерименту було отримано 96 записів для ситуації з повною відсутністю медозбору і 84 запис для ситуації медозбору.

2.9. Попередня обробка даних з четвертого експериментального дослідження

Для попередньої обробки отриманих записів в діапазоні частот від 125 до 1250 Гц використано гребінку із одинадцяти 1/3-октавних нерекурсивних смугових фільтрів Чебишова другого роду [21], так як і в третьому експерименті. Значення центральних частот f_0 та смуг пропускання Δf фільтрів наведено в табл. 6. Порядок фільтрів становив 5000, що дозволило реалізувати амплітудно-частотну характеристику (АЧХ) фільтрів практично прямокутної форми із низьким рівнем бокових пелюстків (-110 дБ).

Шляхом оцінювання дисперсії D_i сигналів із виходів кожного i -го фільтра, було отримано значення G_i , $i = \overline{1,11}$ оцінок спектру потужності сигналів, що аналізуються (1). Для знаходження «первинних» класифікаційних ознак використано значення G_i , $i = \overline{1,11}$ нормовані за по дисперсією D процесу, що аналізується (2).

РОЗДІЛ 3. ОБРОБКА ОТРИМАНИХ ДАНИХ. РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ.

3.1 Обробка даних, отриманих у результаті першого експерименту

Для обробки даних, отриманих у результаті першого експерименту було використано нерекурсивні смугові третьоктавні фільтри Чебишова другого роду [21]. Одинадцять фільтрів, які діяли в діапазоні частот від 125 до 1250Гц. Смуги пропускання фільтрів Δf та середньгеометрична частота $1/3$ октавної смуги f_0 наведені в таблиці 6. Порядок фільтрів становив 5000, що дозволило реалізувати амплітудно-частотну характеристику (АЧХ) фільтрів практично прямокутної форми із низьким рівнем бокових пелюсток. Для прикладу АЧХ одного з фільтрів наведено на рисунку 44. З кожного запису, отриманого в результаті першого експерименту, було отримано одинадцять відфільтрованих сигналів. Було визначено середній рівень звукового тиску кожного з відфільтрованих сигналів:

$$L(f_0) = 20 \cdot \lg(\sigma(f_0)/P_0) \quad (3)$$

$\sigma(f_0)$ - стандартне відхилення відфільтрованого сигналу з частотою $1/3$ октавної смуги f_0 .

Отримані значення рівнів звукового тиску були усереднені по смугам пропускання. Зображення розподілу рівня звукового тиску відносно частоти для кожного з записів (зображені різним кольором), що були отримані в результаті експерименту, та усереднений рівень звукового тиску по смугам пропускання для кожного конкретного вулика представлено на рисунках 45-50.

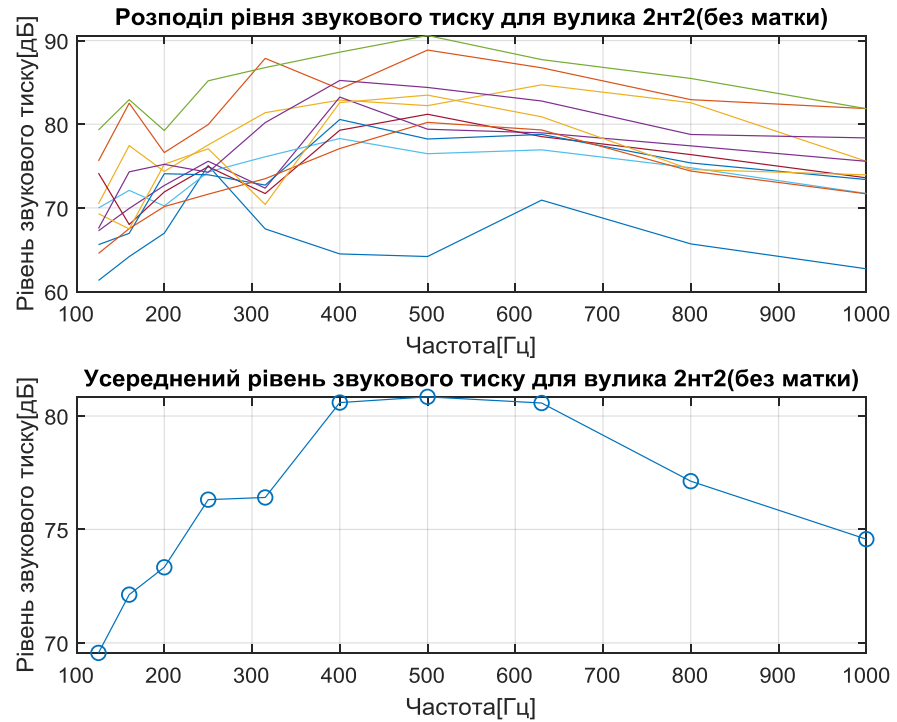


Рис. 45. Розподіл рівня звукового тиску відносно 1/3 октавних частотних смуг для нуклеуса 2нт2, королева відсутня.

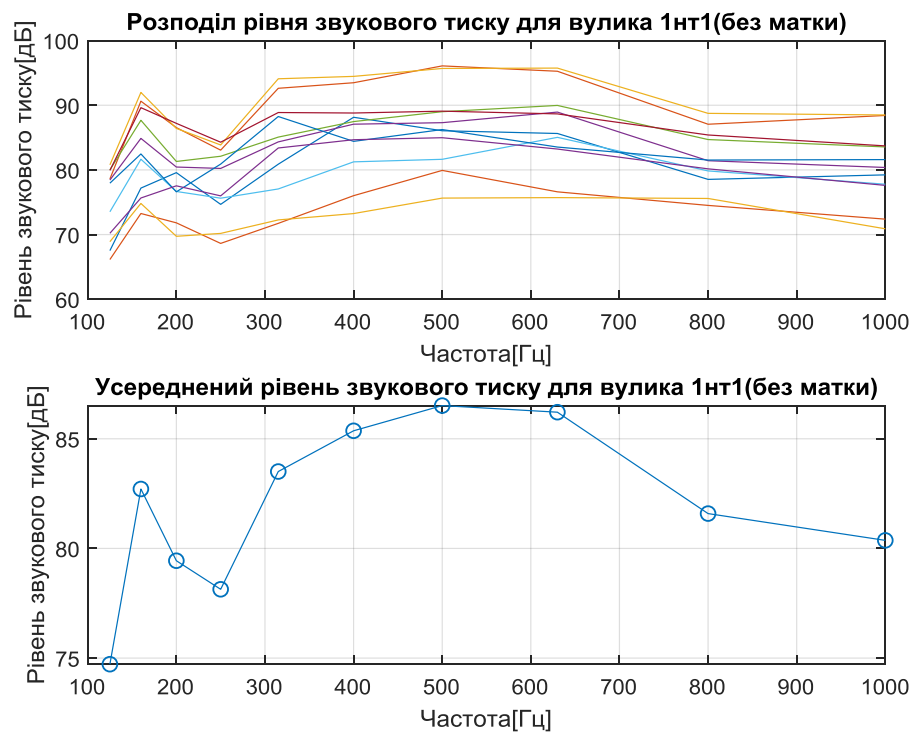


Рис. 46. Розподіл рівня звукового тиску відносно 1/3 октавних частотних смуг для нуклеуса 1нт1, королева відсутня.

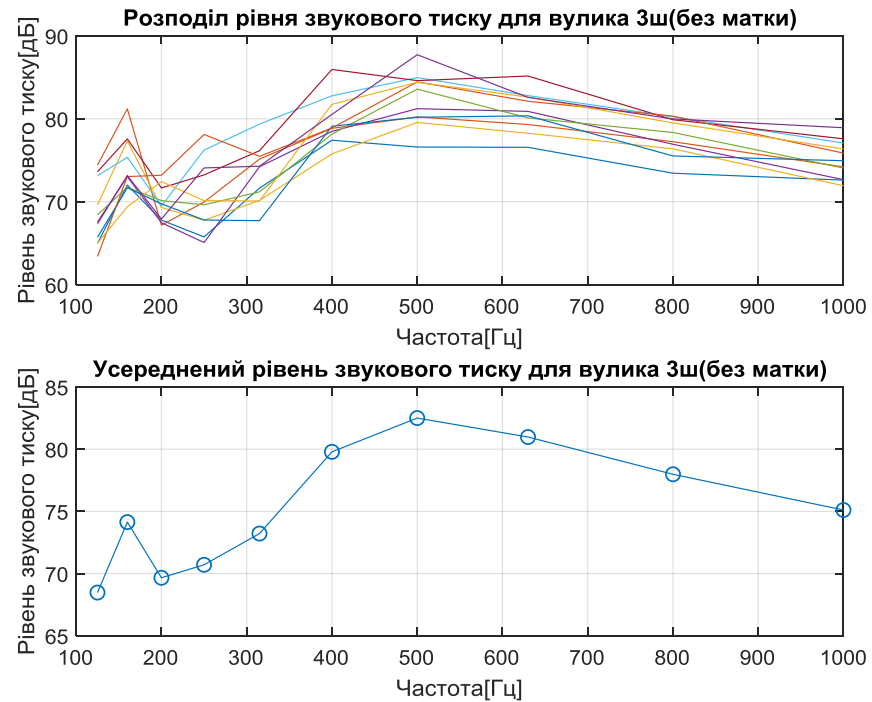


Рис. 47. Розподіл рівня звукового тиску відносно 1/3 октавних частотних смуг для нуклеуса 3ш, королева відсутня.

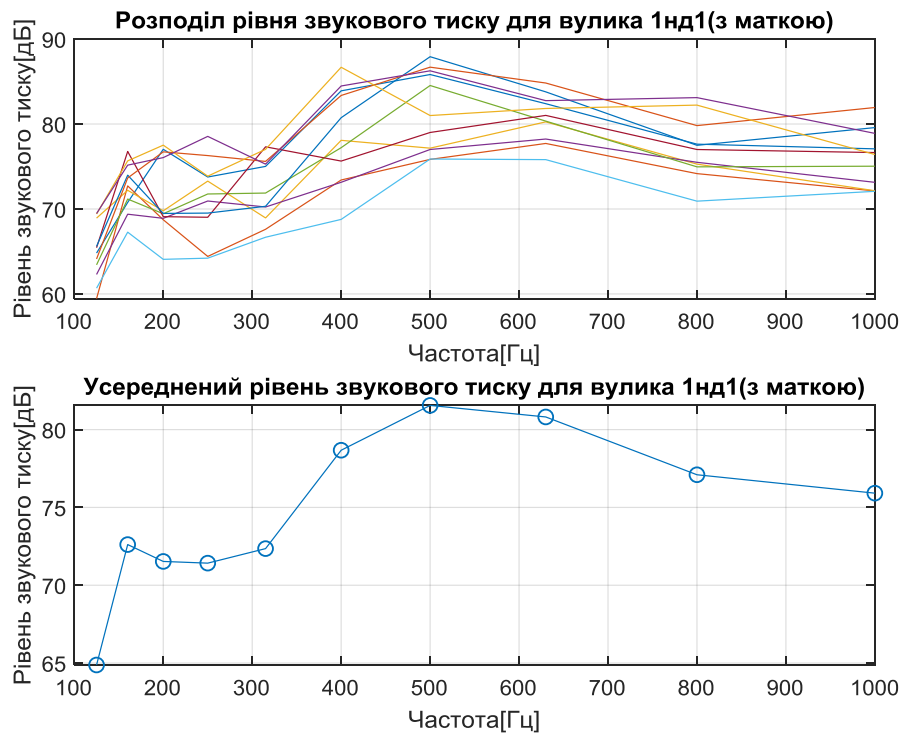


Рис. 48. Розподіл рівня звукового тиску відносно 1/3 октавних частотних смуг для нуклеуса 1нд1, королева присутня.

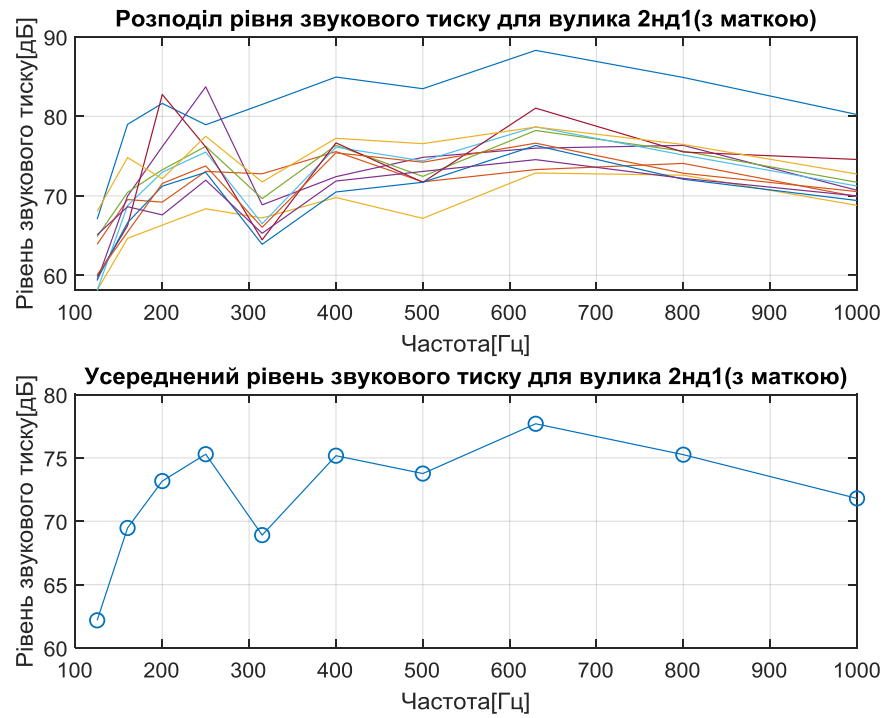


Рис. 49. Розподіл рівня звукового тиску відносно 1/3 октавних частотних смуг для нуклеуса 2нд1, королева присутня.

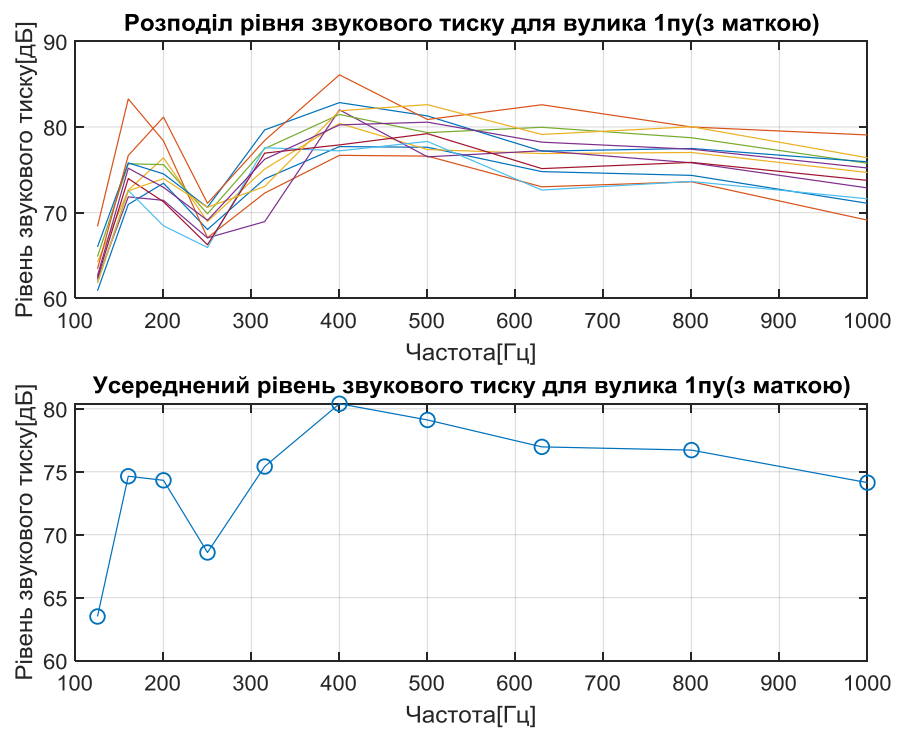


Рис. 50. Розподіл рівня звукового тиску відносно 1/3 октавних частотних смуг для нуклеуса 1пу, королева присутня.

З вищевказаних зображень розподілу видно що, рівень звукового тиску відносно 1/3 октавних частотних смуг для кожного з нуклеусів має значний розкид. Записи для кожного з нуклеусів мають різний розкид, що коливається від ± 5 дБ до ± 15 дБ. Це свідчить про невисоку якість отриманої в результаті експерименту вибірки. З зображень усередненого розподілу випливає що максимум рівнів звукового тиску на частоті 500Гц зберігається для нуклеусів у випадку втрати королеви. На зображеннях усередненого розподілу для нуклеусів у випадку присутності королеви, максимумами рівнів звукового тиску на смугах пропускання зсуваються як в області низьких так і в області середніх частот. Отримані для кожного з нуклеусів усереднені розподіли були знову усереднені (рисунок 51), щоб отримати два зображення розподілу рівнів звукового тиску по смугам пропускання для різних класифікаційних станів.

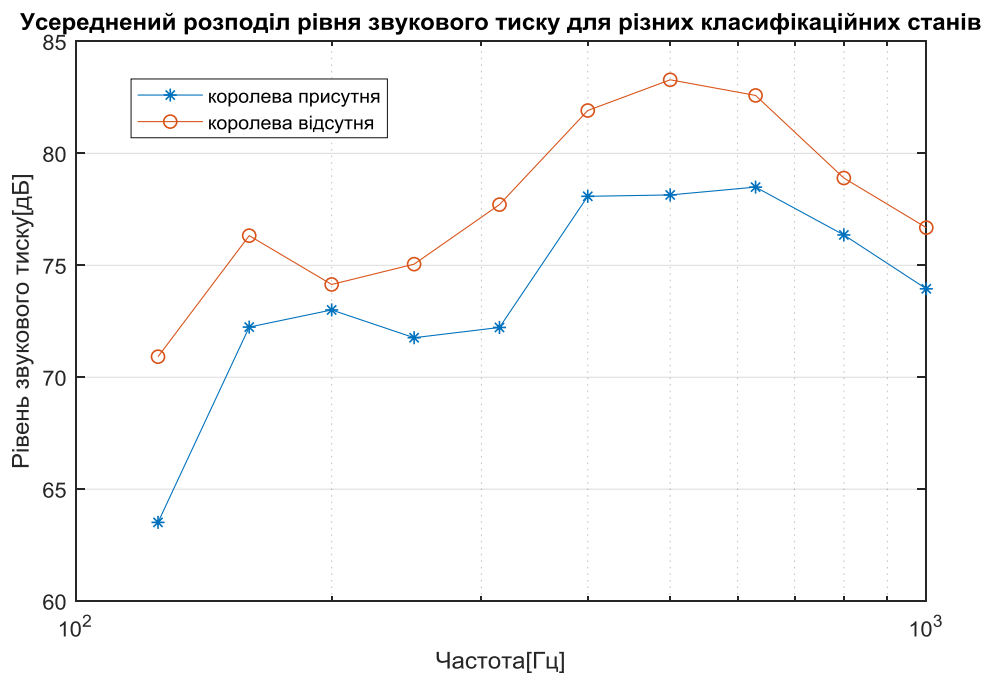


Рис. 51. Усереднений розподіл рівня звукового тиску відносно 1/3 октавних частотних смуг для двох класифікаційних станів.

Розглянувши рисунок 51 помітно, що розподіл рівнів звукового тиску для двох класифікаційних станів відрізняються тим, що на усередненому розподілі

отриманому в ситуації з відсутньою у колонії королевою, локальний максимум рівня звукового тиску знаходиться на частоті 160Гц, а в ситуації з присутньою королевою максимум зміщується на частоту 200Гц. Також в ситуації з відсутньою королевою, локальний максимум присутній на частоті 500Гц, на відміну від ситуації коли королева була присутня в колонії. Слід зазначити що усереднений розподіл в ситуації з відсутньою у колонії королевою має вищий рівень звукового тиску на всіх 1/3 октавних частотних смугах, відносно іншого класу. Отримані спектральні характеристики узгоджуються з результатами досліджень С. Фераррі, Д. Сілва і М. Гуаріно [9], але не дозволяють створити якісну систему класифікації стану сім'ї за обраними класифікаційними ознаками.

3.2. Обробка даних, отриманих у результаті другого експерименту

Для обробки даних, отриманих у результаті другого експерименту було використано нерекурсивні смугові третьоктавні фільтри Чебишова другого роду [21], так як і при обробці даних отриманих у результаті першого експерименту. Одинадцять фільтрів, які діяли в діапазоні частот від 125 до 1250Гц. Порядок фільтрів становив 5000, що дозволило реалізувати АЧХ фільтрів практично прямокутної форми із низьким рівнем бокових пелюсток. З кожного запису, отриманого в результаті другого експерименту, було отримано одинадцять відфільтрованих сигналів. Було визначено середній рівень звукового тиску кожного з відфільтрованих сигналів за формулою (3). Отримані значення рівнів звукового тиску були усереднені по смугам пропускання. Зображення розподілу рівня звукового тиску відносно частоти для кожного з записів (зображені різним кольором), що були отримані в результаті другого експерименту, та усереднений рівень звукового тиску по смугам пропускання для кожного конкретного вулика представлено на рисунках 52-56.

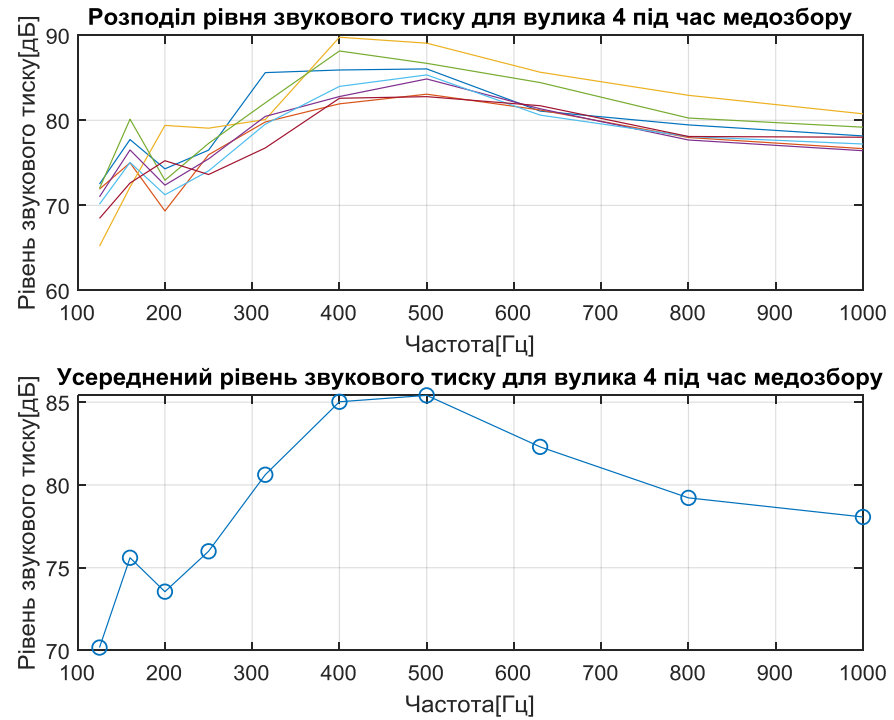


Рис. 52. Розподіл рівня звукового тиску відносно 1/3 октавних частотних смуг для вулика 4 під час медозбору

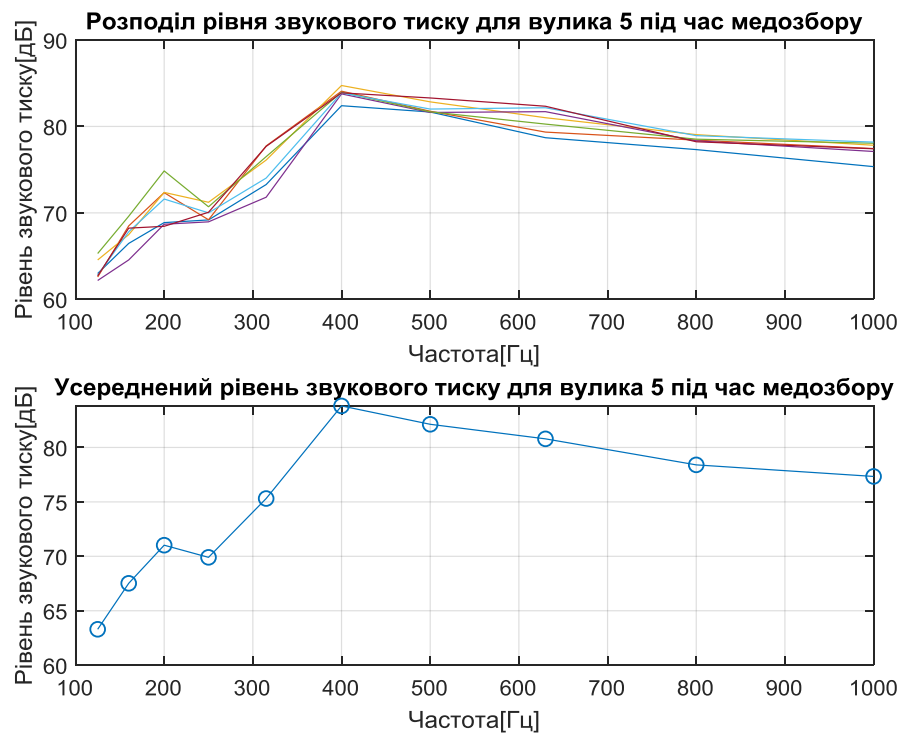


Рис. 53. Розподіл рівня звукового тиску відносно 1/3 октавних частотних смуг для вулика 5 під час медозбору

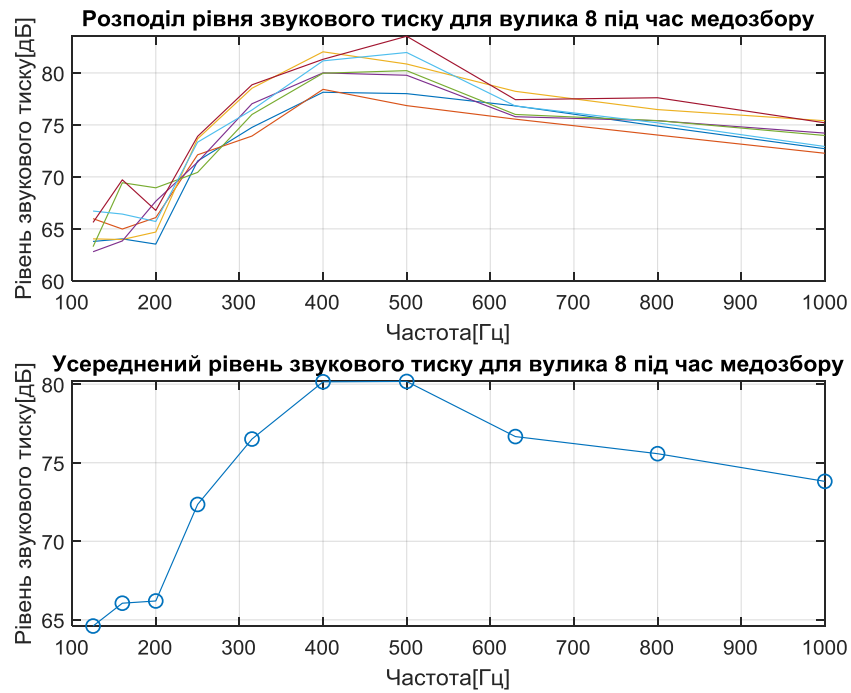


Рис. 54. Розподіл рівня звукового тиску відносно 1/3 октавних частотних смуг для вулика 8 під час медозбору

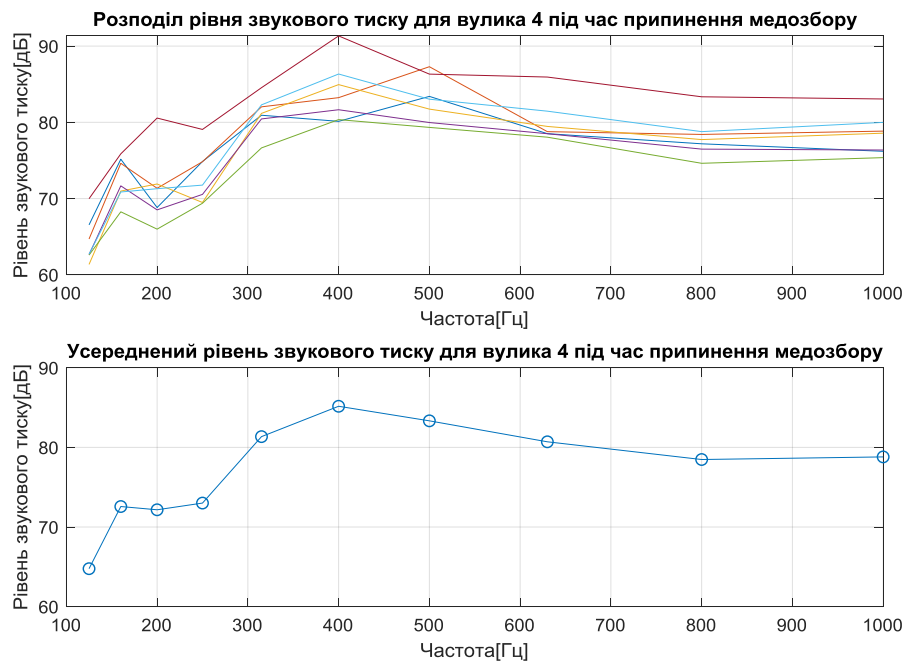


Рис. 55. Розподіл рівня звукового тиску відносно 1/3 октавних частотних смуг для вулика 4 під час припинення медозбору

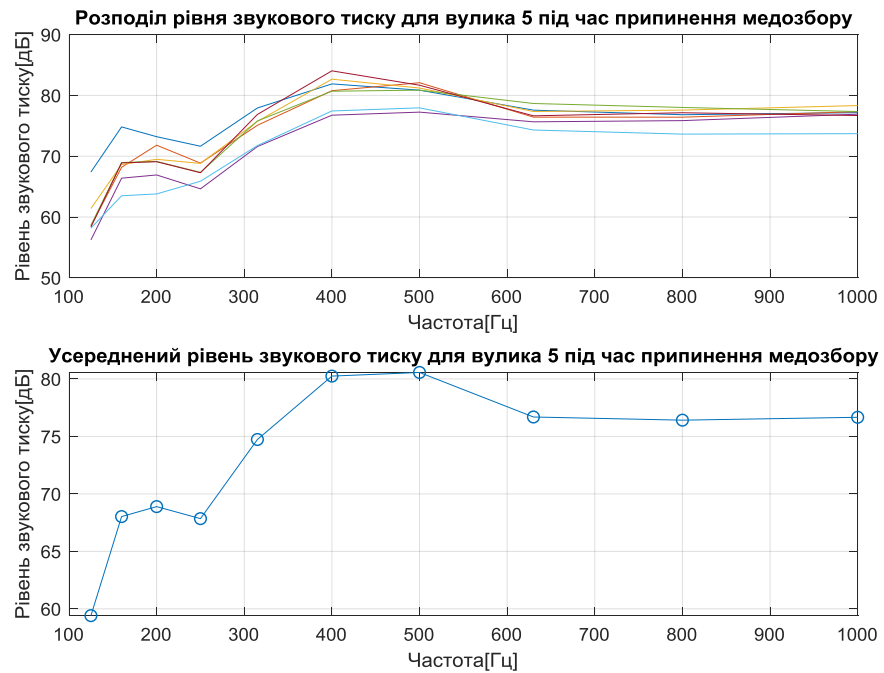


Рис. 56. Розподіл рівня звукового тиску відносно 1/3 октавних частотних смуг для вулика 5 під час припинення медозбору

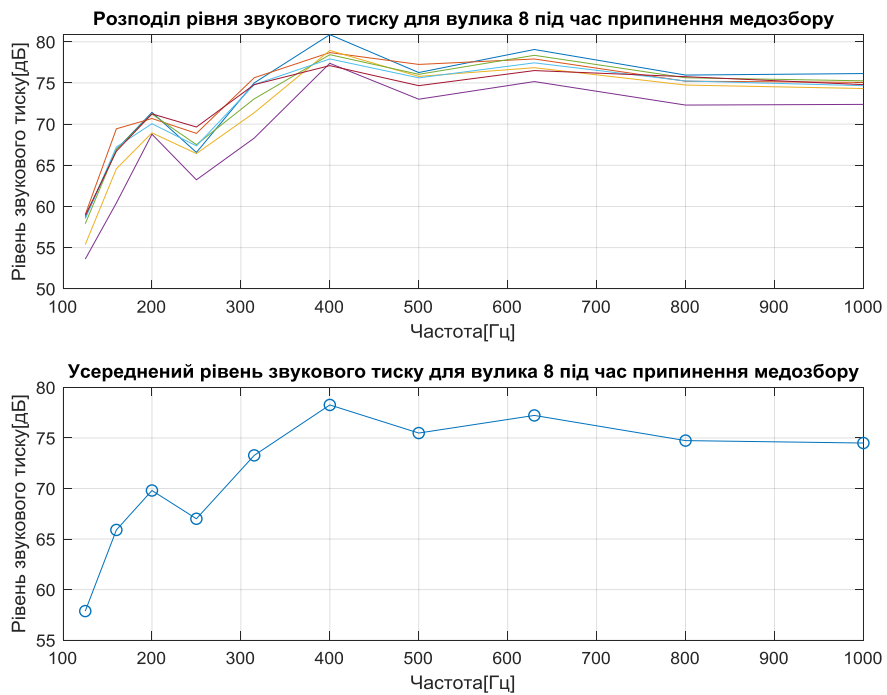


Рис. 57. Розподіл рівня звукового тиску відносно 1/3 октавних частотних смуг для вулика 8 під час припинення медозбору

З вищевказаних зображень розподілу видно що, рівень звукового тиску відносно 1/3 октавних частотних смуг для кожного з нуклеусів має розкид що коливається від ± 5 дБ до ± 10 дБ. На зображеннях усередненого розподілу для всіх вуликів незалежно від класифікаційних станів, присутні максимуми рівнів звукового тиску на смугах частот 400-500Гц і на смугах частот 160-200Гц. Отримані для кожного з вуликів усереднені розподіли були знову усереднені (рисунок 58), щоб отримати два зображення розподілу рівнів звукового тиску по смугам пропускання для різних класифікаційних станів.

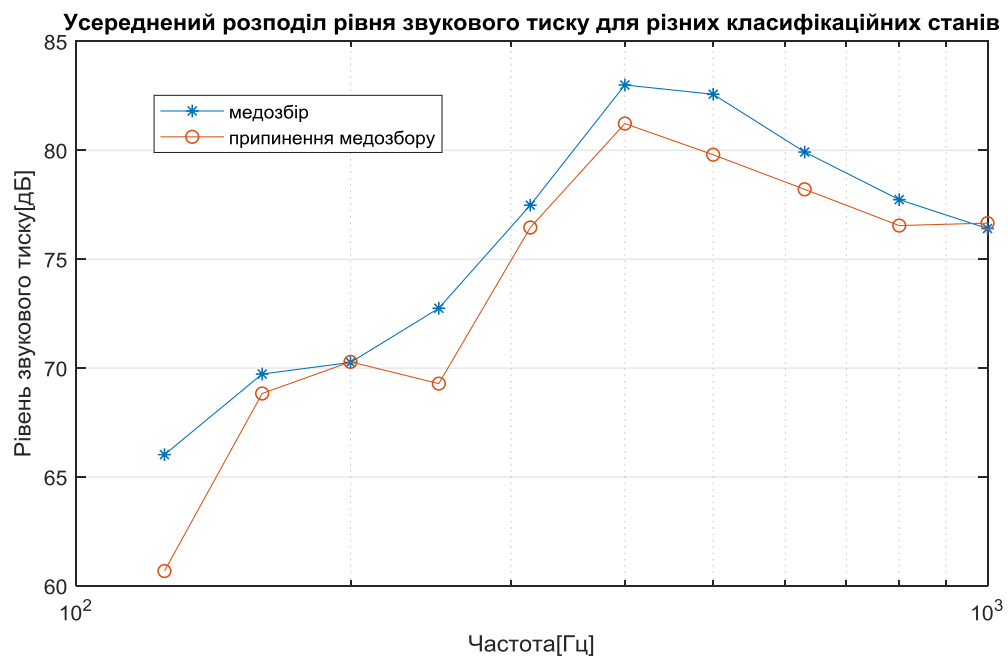


Рис. 58. Усереднений розподіл рівня звукового тиску відносно 1/3 октавних частотних смуг для двох класифікаційних станів.

Розглянувши розподіл рівнів звукового тиску для двох класифікаційних станів на рисунку 58 слід зазначити що міжкласові розподіли відрізняються тільки тим, що на усередненому розподілі, отриманому під час медозбору, рівень звукового тиску на всіх 1/3 октавних частотних смугах, не набагато вищий відносно іншого класу. Локальний максимум рівня звукового тиску, що знаходиться частоті 400Гц, присутній для обох класифікаційних станів.

Отримані спектральні характеристики не можуть бути обраними як класифікаційні ознаки, за якими можна розробити систему класифікації так як присутні у обох класифікаційних станах та суперечать з результатами досліджень С. Фераррі, Д. Сілва і М. Гуаріно [6], та Еськова [7],[8],[9].

3.3 Оцінка інформативності первинних класифікаційних ознак для даних, отриманих у результаті четвертого експерименту.

Після первинної обробки даних, отриманих у результаті четвертого експерименту, була проведена оцінка інформативності первинних класифікаційних ознак. Тобто, шляхом оцінювання дисперсії сигналів із виходів кожного з 11 фільтрів, були отримані оцінки спектру потужності сигналів, що були нормовані за дисперсією процесу. Після чого була знайдена міра відмінності оцінок спектральної густини потужності на 1/3 октавних частотних смугах, тобто оцінка інформативності первинних класифікаційних ознак.

В якості міри інформативності первинних класифікаційних ознак P_i використано так звану «відстань Махаланобіса» [22]:

$$d_i = \frac{|m_{i1} - m_{i2}|}{\sqrt{\sigma_{i1}\sigma_{i2}}}, \quad (4)$$

де m_{i1} та m_{i2} - математичні очікування випадкових величин P_{i1} та P_{i2} , а σ_{i1} та σ_{i2} - стандартні відхилення випадкових величин P_{i1} та P_{i2} . Цифрами «1» та «2» позначено номери класів.

На рисунку 59 зображено значення відстані Махаланобіса для розглянутих значень P_i , $i = \overline{1,11}$.

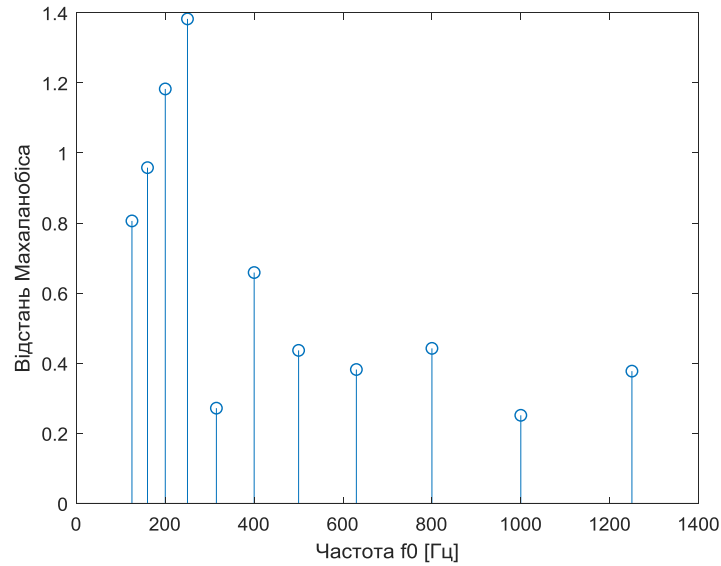


Рис. 59. Значення відстані Махаланобіса для P_i

Згідно із рис. 59, в якості найбільш інформативних ознак доцільно використовувати значення оцінок спектру потужності P_3 ($f_0 = 200$ Гц, $\Delta f = 44$ Гц) та P_4 ($f_0 = 250$ Гц, $\Delta f = 56$ Гц).

3.4. Створення систем класифікації для контролю активності бджіл під час медозбору.

Якщо трактувати ці значення як координати (P_4, P_3) точок на площині, отримуємо два скупчення точок. На рис. 60 точки, що відносяться до класу 1, позначені кружечками, а до класу 2 - ромбиками. До класу 1 відносяться значення P_3 та P_4 що були отримані з записів під час медозбору, до класу 2 відносяться відповідні значення отримані під час повного припинення медозбору.

Рівняння кривої $P_3 = f(P_4)$, що розділяє ці два скупчення точок, визначає правило прийняття рішення:

$$\begin{aligned} &\text{клас 1, якщо } P_3 \geq f(P_4), \\ &\text{клас 2, якщо } P_3 < f(P_4). \end{aligned} \tag{5}$$

Зображену на рисунку 60 поділяючу криву було побудовано на основі даних навчальної вибірки за методом мінімальної середньоквадратичної похибки.

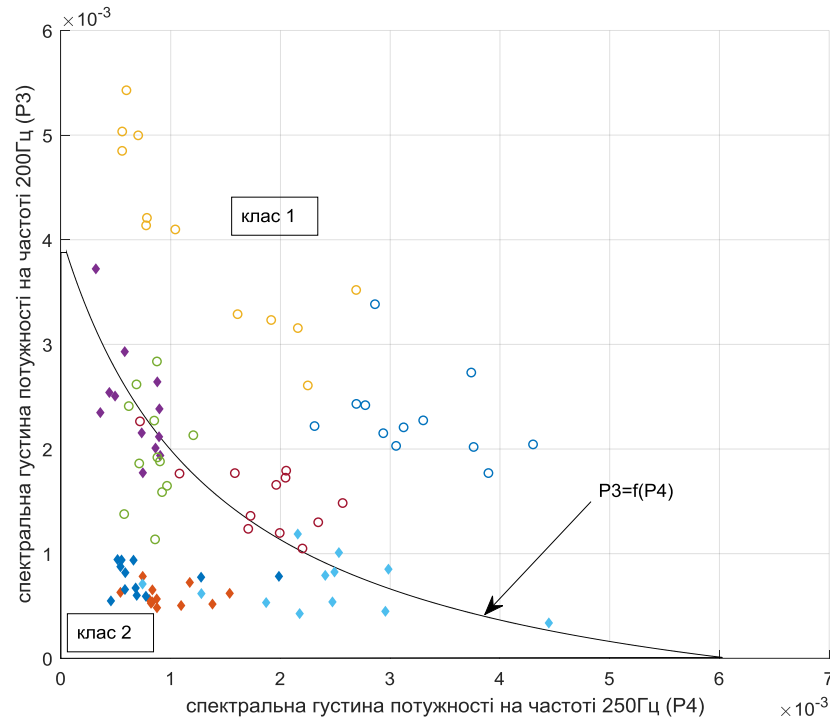


Рис. 60. Значення пари ознак $\{P_3, P_4\}$ для класу 1 (кружечки) та класу 2 (ромбики)

Рівняння поділяючої кривої:

$$P_3 = \frac{0.01}{1400 \cdot P_4 + 2} - 9.5 \cdot 10^{-4}. \quad (6)$$

Оцінка ймовірності правильної класифікації за парою ознак $\{P_3, P_4\}$ для навчальної вибірки даних становила $P_{\text{клн}} = 0,81$.

Оскільки навчальна вибірка містила записи по чотирьом вуликам для обох класів, кружечки та ромбики відображено різними кольорами. Досить цікавим є те, що скупчення кружечків та ромбиків різних кольорів рознесені на площині. Це свідчить про відмінність характеру шуму не тільки для різних

класів, але й навіть для різних вуликів, що в подальшому може бути використано для класифікації бджолиних сімей.

Оцінка ймовірності правильної класифікації за парою ознак $\{P_3, P_4\}$ для тестової вибірки даних, яка складалася із записів для трьох вуликів першого класу і чотирьох вуликів другого класу, становила $R_{клт} = 0,96$.

Для виявлення можливості спрощення системи класифікації, доцільно розглянуто випадок, коли використовується лише одна класифікаційна ознака, P_3 або P_4 .

На рисунку 61 представлено оцінки умовних розподілів щільності ймовірності $W(P_3 | кл1)$ та $W(P_3 | кл2)$ випадкової величини P_4 , а на рисунку 62 представлено аналогічні оцінки розподілів $W(P_4 | кл1)$ та $W(P_4 | кл2)$ для величини P_4 . Як видно з рис. 61, для класифікаційної ознаки P_3 , оцінки щільності ймовірності першого та другого класів суттєво перекриваються, що утруднює використання простого правила прийняття рішення:

$$\begin{aligned} &\text{клас 1, якщо } P_3 \geq \Pi, \\ &\text{клас 2, якщо } P_3 < \Pi, \end{aligned} \quad (7)$$

де Π - певне порогове значення випадкової величини P_3 . Тому використаємо байесівське правило прийняття рішень:

$$\begin{aligned} &\text{клас 1, якщо } P_{кл1} \cdot W(P_3 | кл1) \geq P_{кл2} \cdot W(P_3 | кл2), \\ &\text{інакше клас 2,} \end{aligned} \quad (8)$$

де $P_{кл1}$ та $P_{кл2}$ - апіорні ймовірності 1-го та 2-го класів, відповідно. Оскільки апіорні ймовірності $P_{кл1}$ та $P_{кл2}$ є невідомими, приймемо $P_{кл1} = P_{кл2} = 0,5$.

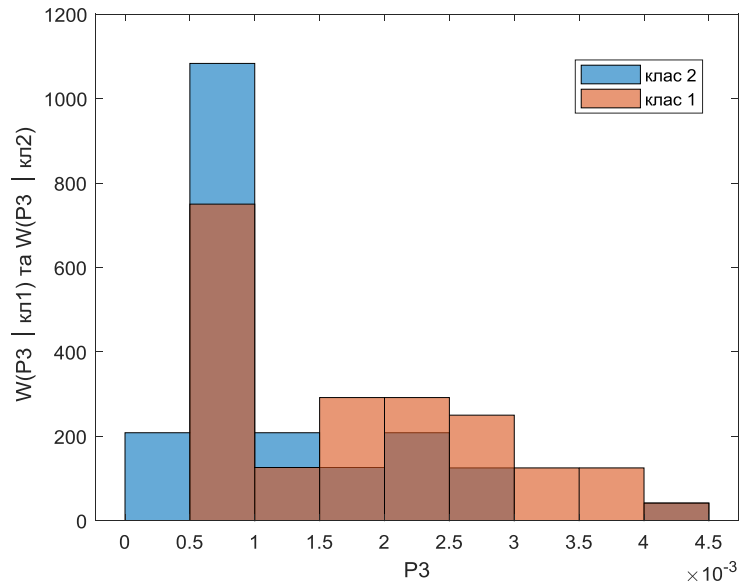


Рис. 61. Оцінки умовних розподілів $W(P_3 | кл1)$ та $W(P_3 | кл2)$

Обчислення показали, що для класифікаційної ознаки P_3 оцінка ймовірності правильної класифікації для навчальної вибірки становить $R_{клн} = 0,66$, а для тестової вибірки $R_{клт} = 0,90$.

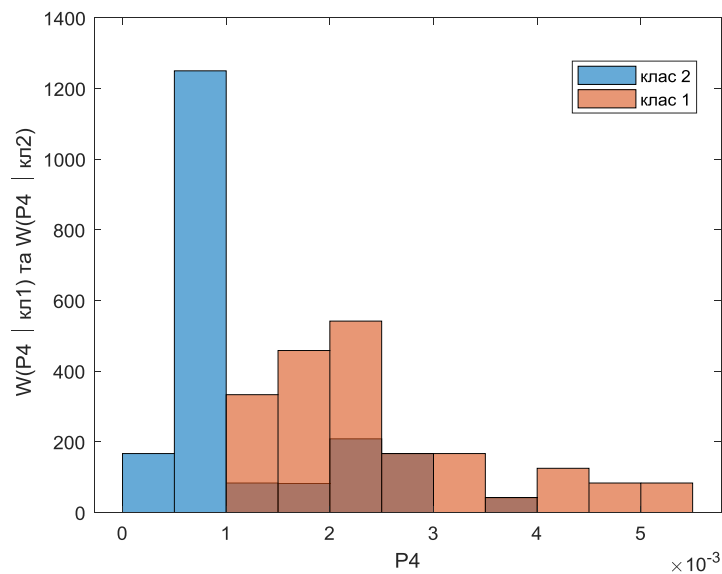


Рис. 62. Оцінки умовних розподілів $W(P_4 | кл1)$ та $W(P_4 | кл2)$

Для класифікаційної ознаки P_4 (рис. 62) оцінки ймовірності правильної класифікації становлять $R_{клн} = 0,86$ та $R_{клт} = 0,84$ для навчальної та тестової вибірок, відповідно.

Також були розглянуті системи класифікації при нормуванні значень оцінок спектральної густини потужності по сумі спектральних потужностей сигналів відфільтрованих на інформативних частотних смугах. Для класифікації по двом інформаційним ознакам при такому методі нормування обидві ознаки пов'язані лінійною залежністю: $P_1 + P_2 = 1 \Rightarrow P_1 = 1 - P_2$, що робить класифікацію по двом ознакам недоцільною так як всі точки на площині шикуються у пряму. Значення P_1 і P_2 , відповідно, є нормованими по сумі потужностей відфільтрованих сигналів значеннями оцінок спектральної густини потужності. Для систем класифікації записів з використанням окремо кожної з класифікаційних ознак на рисунку 63 представлено оцінки умовних розподілів щільності ймовірності $W(P_1 | кл1)$ та $W(P_1 | кл2)$ випадкової величини P_1 , а на рисунку 64 представлено аналогічні оцінки розподілів $W(P_2 | кл1)$ та $W(P_2 | кл2)$ для величини P_2 . Навчальна і тестова вибірка для даної класифікації еквівалентна навчальній і тестовій вибірці у попередніх системах класифікації.

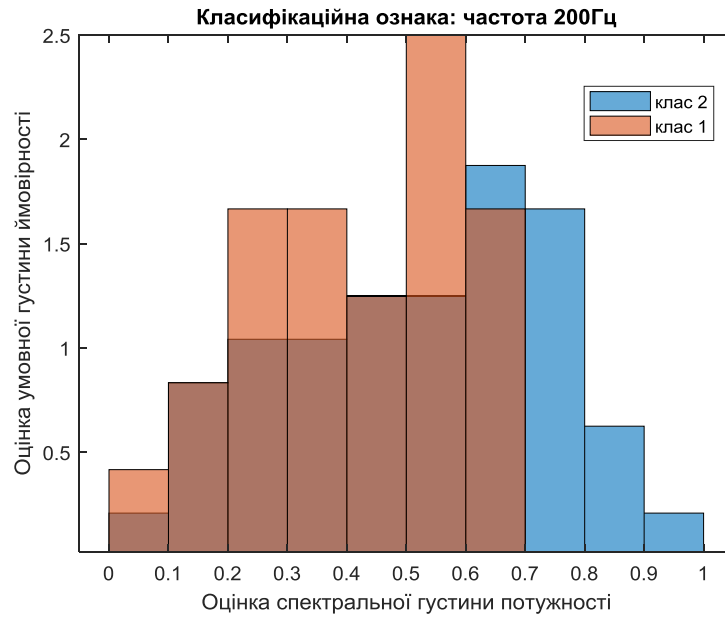


Рис. 63. Оцінки умовних розподілів $W(P_1 | кл1)$ та $W(P_1 | кл2)$

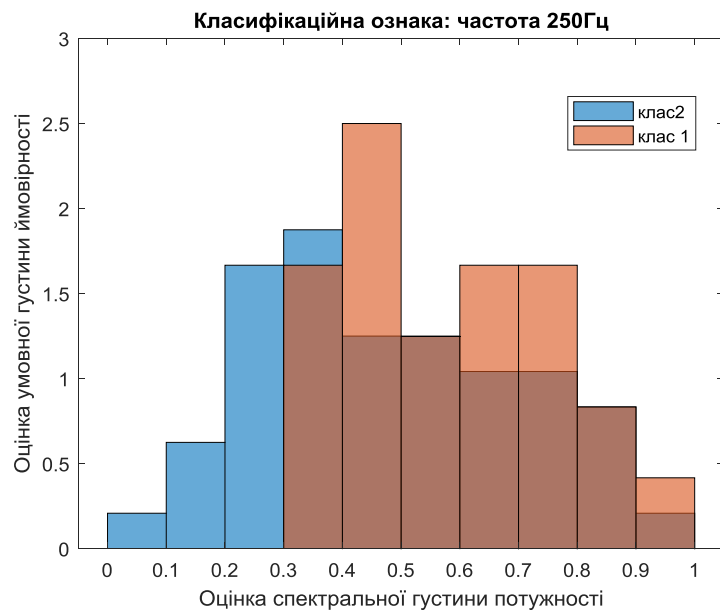


Рис. 64. Оцінки умовних розподілів $W(P_2 | кл1)$ та $W(P_2 | кл2)$

Як видно з рисунку 63 та рисунку 64 для обох класифікаційних ознак гістограми першого та другого класів дуже суттєво перекривають одне одного. Дані гістограми являють собою якісну характеристику класифікації і вказують на те, що ймовірність правильної класифікації має бути суттєво гірше ніж у попередніх способах класифікації. Щоб підтвердити це кількісно, було

використано байесівське і правило прийняття рішень (8) що було розглянуто у вищеописаній системі класифікації. При використанні однієї класифікаційної ознаки, а саме середньгеометричної частоти 250Гц або середньгеометричної частоти 200Гц відповідно, отримані наступні результати: для класифікаційної ознаки P_1 (рис. 63) оцінка ймовірності правильної класифікації для навчальної вибірки становить $R_{клн} = 0,65$, а для тестової вибірки $R_{клт} = 0,22$.

Для класифікаційної ознаки P_2 (рис. 64) оцінки ймовірності правильної класифікації становлять $R_{клн} = 0,66$ та $R_{клт} = 0,2$ для навчальної та тестової вибірок, відповідно. Отже системи класифікації які використовують значення P_1 і P_2 , тобто нормовані по сумі потужностей відфільтрованих сигналів, а не по дисперсії процесу, значення оцінок спектральної густини потужності, являються в даному випадку не ефективними та не будуть розглянуті в подальшому.

3.5. Порівняння ефективності систем класифікації для контролю активності бджіл під час медозбору

Щоб порівняти між собою ефективні системи класифікації стану бджолої сім'ї під час медозбору з використанням пари ознак $\{P_3, P_4\}$ або однієї ознаки P_3 або P_4 , зручно представити відповідні значення $R_{клн}$ та $R_{клт}$, а також середнього значення $R_{сер} = (R_{клн} + R_{клт})/2$ у вигляді таблиці (табл. 7) та графіку (рис. 65).

Табл. 7. Оцінки ймовірностей правильної класифікації

	$\{P_3, P_4\}$	P_3	P_4
$R_{клн}$	0,81	0,66	0,86
$R_{клт}$	0,96	0,90	0,84
$R_{сер}=(R_{клн}+R_{клт})/2$	0,885	0,78	0,85

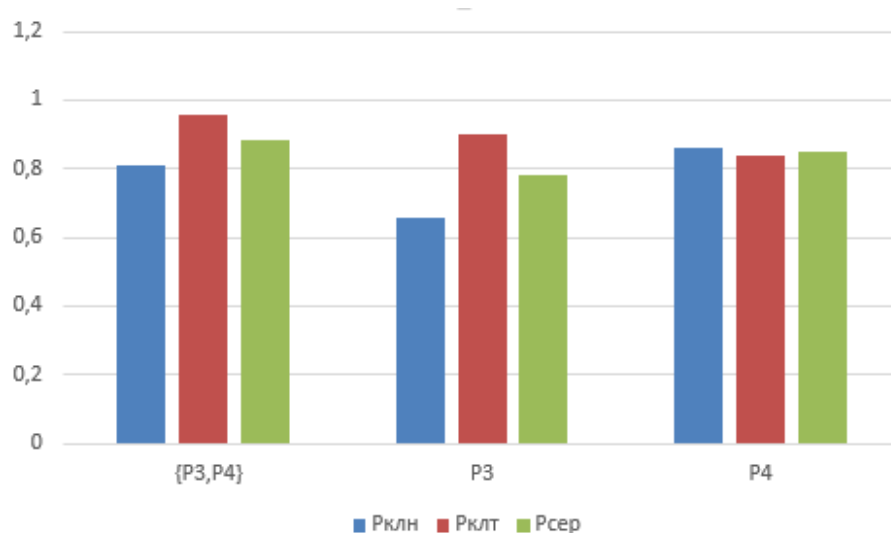


Рис. 65. Оцінки ймовірності правильної класифікації

Отримані результати свідчать, що ефективність (за критерієм ймовірності правильної класифікації) пари класифікаційних ознак $\{P_3, P_4\}$ є дещо вищою порівняно із ефективністю однієї ознаки P_3 або P_4 . Якщо порівнювати між собою ознаки P_3 або P_4 , то ознака P_4 виявилася дещо кращою, хоча різниця в якості класифікації не дуже велика.

Зроблені висновки щодо інформативності певних частотних ділянок спектру сигналу узгоджуються із результатами праць [7] та [9], що свідчить про достовірність отриманих в даній роботі результатів.

3.6. Оцінка інформативності первинних класифікаційних ознак для даних, отриманих у результаті третього експерименту.

Після первинної обробки даних, отриманих у результаті третього експерименту, була створена вибірка що складалася з оцінок спектру потужності сигналів, що були нормовані за дисперсією процесу, для одинадцяти вуликів з вилученою королевою (експеримент 3) та для одинадцяти вуликів з присутньою королевою (експеримент 4). Слід зазначити що загальна вибірка була поділена на навчальну та тестову. Навчальна вибірка містить записи по шести вуликах для обох класів. Тестова вибірка містить

записи по п'яти вуликах для обох класів. По загальній вибірці була проведена оцінка інформативності первинних класифікаційних ознак, тобто була знайдена міра відмінності оцінок спектральної густини потужності на 1/3 октавних частотних смугах.

В якості міри інформативності первинних класифікаційних ознак P_i використано так звану «відстань Махаланобіса» [22], формула (4).

На рисунку 66 зображено значення відстані Махаланобіса для розглянутих значень P_i , $i = \overline{1,11}$.

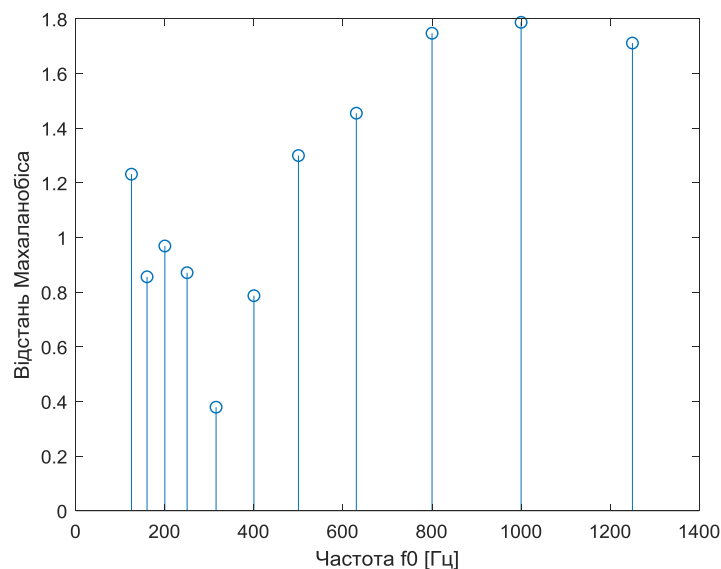


Рис. 66. Значення відстані Махаланобіса для P_i

Згідно із рисунку 66, в якості найбільш інформативних ознак доцільно використовувати значення оцінок спектру потужності P_{10} ($f_0 = 800$ Гц, $\Delta f = 190$ Гц) та P_{11} ($f_0 = 1000$ Гц, $\Delta f = 220$ Гц).

3.7. Створення систем класифікації для виявлення втрати бджолоїної королеви у колонії

Якщо трактувати ці значення як координати (P_{10}, P_{11}) точок на площині, отримуємо два скупчення точок. На рис. 67 точки, що відносяться до класу 1, позначені кружечками, а до класу 2 - ромбиками. Слід нагадати що до класу 1

відносяться значення P_{10} та P_{11} які були отримані з бджолиних колоній в яких королева була вилучена, до класу 2 відносяться відповідні значення які були отримані з бджолиних колоній в яких королева була присутня.

Рівняння кривої $P_{10} = f(P_{11})$, що розділяє ці два скупчення точок, визначає правило прийняття рішення:

$$\begin{aligned} &\text{клас 1, якщо } P_{10} \geq f(P_{11}), \\ &\text{клас 2, якщо } P_{10} < f(P_{11}). \end{aligned} \quad (9)$$

Рівняння поділяючої кривої:

$$P_{10} = \frac{3 \cdot 10^{-5}}{9000 \cdot P_{11}}. \quad (10)$$

Зображену на рисунку 67 поділяючу криву було побудовано на основі даних навчальної вибірки за методом мінімальної середньоквадратичної похибки.

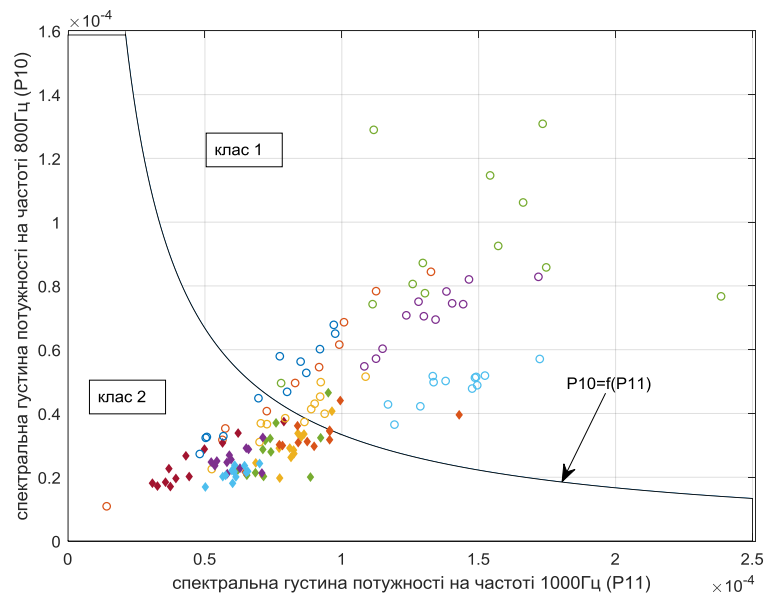


Рис. 67. Значення пари ознак $\{P_{10}, P_{11}\}$ для класу 1 (кружечки) та класу 2 (ромбики)

Оцінка ймовірності правильної класифікації за парою ознак $\{P_{10}, P_{11}\}$ для навчальної вибірки даних становила $R_{\text{клн}} = 0,87$. Оцінка ймовірності правильної класифікації для тестової вибірки даних, становила $R_{\text{клт}} = 0,84$.

Для виявлення можливості спрощення системи класифікації, доцільно розглянуто випадок, коли використовується лише одна класифікаційна ознака, P_{10} або P_{11} . На рисунку 68 представлено оцінки умовних розподілів щільності ймовірності $W(P_{10} | \text{кл1})$ та $W(P_{10} | \text{кл2})$ випадкової величини P_{10} .

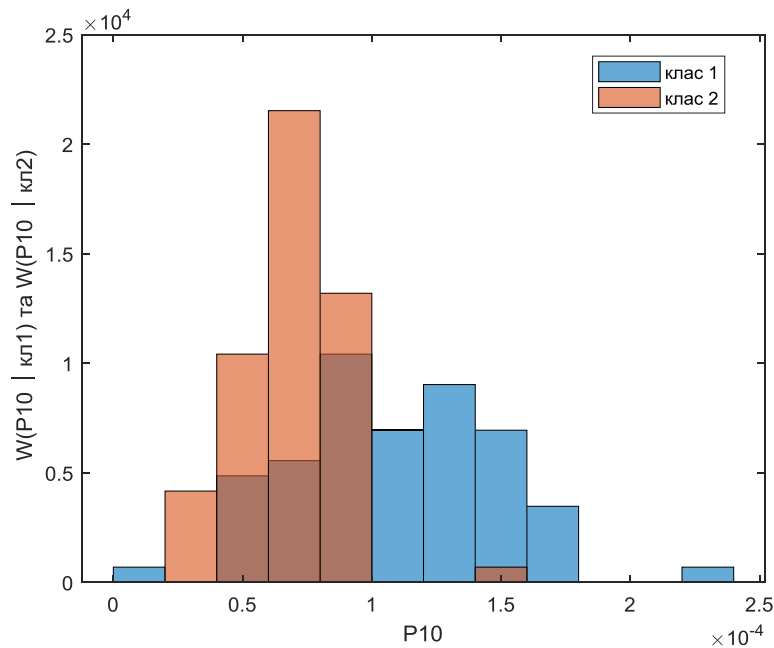


Рис. 68. Оцінки умовних розподілів $W(P_{10} | \text{кл1})$ та $W(P_{10} | \text{кл2})$

Як видно з рис. 68, для класифікаційної ознаки P_{10} , оцінки щільності ймовірності першого та другого класів дещо перекриваються, що утруднює використання простого правила прийняття рішення:

$$\begin{aligned} &\text{клас 1, якщо } P_{10} \geq \Pi, \\ &\text{клас 2, якщо } P_{10} < \Pi, \end{aligned} \quad (11)$$

де Π - певне порогове значення випадкової величини P_{10} . Тому використаємо байєсівське правило прийняття рішень:

$$\begin{aligned} &\text{клас 1, якщо } P_{\kappa 1} \cdot W(P_{10} | \kappa 1) \geq P_{\kappa 2} \cdot W(P_{10} | \kappa 2), \\ &\text{інакше клас 2,} \end{aligned} \quad (12)$$

де $P_{\kappa 1}$ та $P_{\kappa 2}$ - апіорні ймовірності 1-го та 2-го класів, відповідно. Оскільки апіорні ймовірності $P_{\kappa 1}$ та $P_{\kappa 2}$ є невідомими, прийнемо $P_{\kappa 1} = P_{\kappa 2} = 0,5$.

Обчислення показали, що для класифікаційної ознаки P_{10} оцінка ймовірності правильної класифікації для навчальної вибірки становить $R_{\kappa 1} = 0,78$, а для тестової вибірки $R_{\kappa 2} = 0,62$.

На рисунку 69 представлено оцінки умовних розподілів щільності ймовірності $W(P_{11} | \kappa 1)$ та $W(P_{11} | \kappa 2)$ для величини P_{11} .

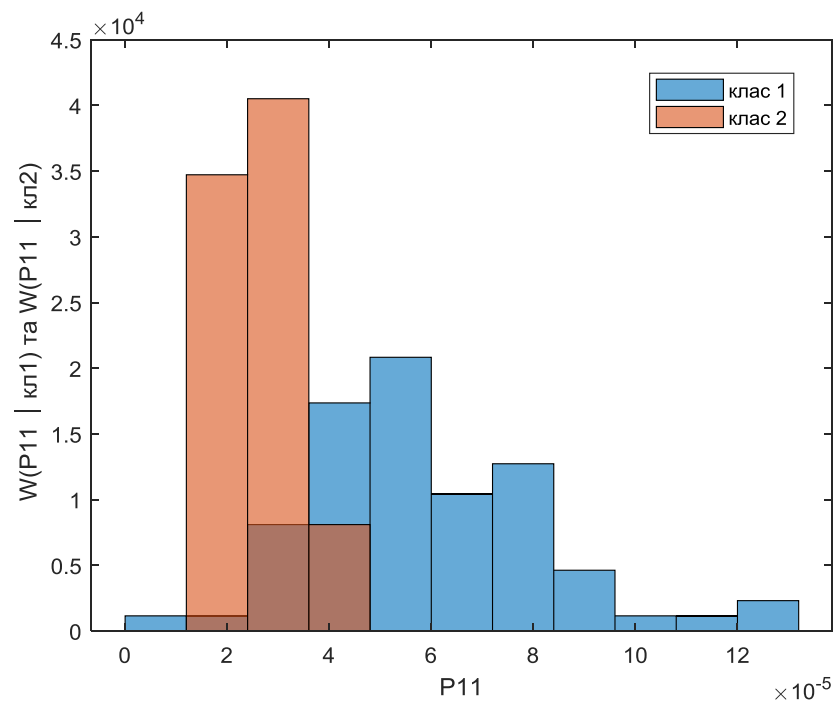


Рис. 69. Оцінки умовних розподілів $W(P_{11} | \kappa 1)$ та $W(P_{11} | \kappa 2)$.

Для класифікаційної ознаки P_{11} , використаємо байесівське правило прийняття рішень:

$$\begin{aligned} &\text{клас 1, якщо } P_{\kappa l1} \cdot W(P_{11} | \kappa l1) \geq P_{\kappa l2} \cdot W(P_{11} | \kappa l2), \\ &\text{інакше клас 2,} \end{aligned} \quad (13)$$

де $P_{\kappa l1}$ та $P_{\kappa l2}$ - апіорні ймовірності 1-го та 2-го класів, відповідно. Оскільки апіорні ймовірності $P_{\kappa l1}$ та $P_{\kappa l2}$ є невідомими, прийmemo $P_{\kappa l1} = P_{\kappa l2} = 0,5$.

Для класифікаційної ознаки P_{11} (рис. 69) оцінки ймовірності правильної класифікації становлять $R_{\kappa l1} = 0,88$ та $R_{\kappa l2} = 0,71$ для навчальної та тестової вибірок, відповідно.

3.8. Порівняння ефективності систем класифікації для виявлення втрати бджолиної королеви у колонії

Щоб порівняти між собою ефективні системи класифікації стану бджолиної сім'ї у випадку втрати королеви з використанням пари ознак $\{P_{10}, P_{11}\}$ або однієї ознаки P_{10} або P_{11} , зручно представити відповідні значення $R_{\kappa l1}$ та $R_{\kappa l2}$, а також середнього значення $R_{\text{сер}} = (R_{\kappa l1} + R_{\kappa l2})/2$ у вигляді таблиці (табл. 8) та графіку (рис. 70).

Табл. 8. Оцінки ймовірностей правильної класифікації

	$\{P_{10}, P_{11}\}$	P_{10}	P_{11}
$R_{\kappa l1}$	0,87	0,78	0,88
$R_{\kappa l2}$	0,84	0,62	0,71
$R_{\text{сер}} = (R_{\kappa l1} + R_{\kappa l2})/2$	0,855	0,7	0,795

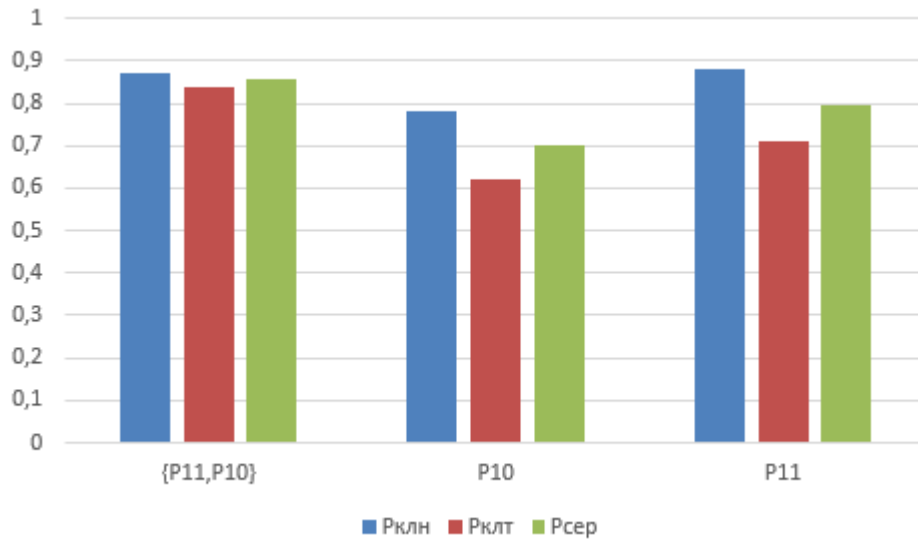


Рис. 70. Оцінки ймовірності правильної класифікації

Отримані результати свідчать, що ефективність (за критерієм ймовірності правильної класифікації) пари класифікаційних ознак $\{P_{10}, P_{11}\}$ є дещо вищою порівняно із ефективністю однієї ознаки P_{10} або P_{11} . Якщо порівнювати між собою ознаки P_{10} або P_{11} , то ознака P_{11} виявилася дещо кращою, хоча різниця в якості класифікації не дуже велика.

ВИСНОВОК

Досліджено можливості побудови простої, з точки зору технічної реалізації, системи моніторингу стану бджолоїної сім'ї за спектральними характеристиками її шуму всередині вулика. Виконано експериментальні дослідження, спрямовані на виявлення ефективних класифікаційних ознак, котрі дозволили б класифікувати стан втрати бджолоїної королеви або характеризували активність під час медозбору у бджолоїній сім'ї. Із використанням гребінки із одинадцяти 1/3-октавних фільтрів виконано експериментальні дослідження оцінок спектральної потужності шуму бджіл в частотному діапазоні 125-1250 Гц. Отримано оцінки інформативності значень спектральної потужності та виявлено, що для класифікації стану медозбору найбільш інформативними класифікаційними ознаками є оцінки спектру потужності шуму бджолоїної сім'ї в околицях частот 200 Гц та 250 Гц. Для класифікації стану втрати королеви найбільш інформативними класифікаційними ознаками є оцінки спектру потужності шуму бджолоїної сім'ї в околицях частот 800 Гц та 1 кГц. Показана доречність нормування значень спектральної потужності шуму по дисперсії шуму, що дозволяє домогтися інваріантності результатів класифікації до рівня шуму. Створено дві системи класифікації для кожного з розглянутих станів. В першій системі класифікації в якості класифікаційних ознак використано два значення оцінок спектру потужності на двох смугах інформативних частот, а в якості правила прийняття рішення використано рівняння граничної кривої, що розділяє скупчення точок конкуруючих класів на площі. В альтернативній системі класифікації в ролі класифікаційної ознаки використано оцінку спектру потужності на одній з інформативних частотних смуг, а в якості правила прийняття рішення використано байесівську стратегію. Тестування працездатності та ефективності створених систем класифікації виконано шляхом оцінювання ймовірності правильної класифікації на навчальних та тестових вибірках даних. Таке

тестування показало, що система класифікації, що визначає наявність або відсутність медозбору, у випадку використання двох класифікаційних ознаки, забезпечує ймовірність правильної класифікації 0,81 для навчальної вибірки даних, та 0,96 для тестової вибірки даних. Аналогічне тестування було проведено для другого різновиду системи класифікації вищевказаного стану бджолої сім'ї, де класифікаційною ознакою слугує лише одне значення спектру. Для випадку, коли це значення оцінювалося на частоті 200 Гц, ймовірність правильної класифікації склала 0,66 для навчальної вибірки даних та 0,90 для тестової вибірки даних. У випадку частоти 250 Гц, ймовірність правильної класифікації становила 0,86 для навчальної вибірки даних та 0,84 для тестової вибірки даних.

Тестування ефективності системи класифікації, що визначає втрату бджолої королеви у колонії, у випадку використання двох класифікаційних ознаки, забезпечує ймовірність правильної класифікації 0,87 для навчальної вибірки даних, та 0,84 для тестової вибірки даних. Аналогічне тестування було проведено для другого різновиду системи класифікації вищевказаного стану бджолої сім'ї, де класифікаційною ознакою слугує лише одне значення спектральної густини потужності. Для випадку, коли це значення оцінювалося на частоті 800 Гц, ймовірність правильної класифікації склала 0,78 для навчальної вибірки даних та 0,62 для тестової вибірки даних. У випадку частоти 1 кГц, ймовірність правильної класифікації становила 0,88 для навчальної вибірки даних та 0,71 для тестової вибірки даних.

Показано, що найбільш точними є системи класифікації, де в якості класифікаційних ознак використано два значення спектральної потужності шуму. Альтернативні системи класифікації, в яких використовується лише одне значення спектру потужності, є простішими для практичної реалізації, проте вони поступаються за точністю класифікації. На основі отриманих

систем класифікації можливо створити автоматизовану систему моніторингу для інформування бджоляра про характер активності бджіл у вуликах.

СПИСОК ЛІТЕРАТУРИ

1. Honeybee varroa vibration and ccd. [Online]. Available:
<http://modernsurvivalblog.com/natural-disasters/the-honeybeevarroa-vibration-and-ccd/> (document)
2. Camazine, S., Visscher, P.K., 1999. House haunting by honey bee swarms: collective decisions and individual behaviour. *Insects Soc.* 46, 348–360.
3. D. Sammataro, U. Gerson, and G. Needham, “Parasitic mites of honey bees: life history, implications, and impact,” *Annual review of entomology*, vol. 45, no. 1, pp. 519–548, 2000.
4. E. F. Woods, Electronic Prediction of Swarming in Bees // *NATURE*, Sep. 19, 1959, vol. 184, P. 842-844.
5. JINDRA J. Využití akustické technologie pro vyhodnocování stavu včelstev. http://jjvcela.sweb.cz/Poslouchame_vcely.html, 2016.
6. D. G. Dietlein, A Method for Remote Monitoring of Activity of Honeybee Colonies by Sound Analysis, *Journal of Apicultural Research*, 24:3, 176-183, 1985. DOI:10.1080/00218839.1985.11100668
7. Єськов Е.К. Структура звукового фону бджолої сім'ї // *Зоол. журн.* 1972. Т. 51, № 7. С. 1018-1024.
8. Єськов Е.К. Екологія медоносної бджоли // *Рязань: Російське слово*, 1995. С. 392
9. Ferrari S. Monitoring of swarming sounds for early detection of the swarming period / S. Ferrari, D. Silva, M. Guarino, D. Berckmans // *Computers and Electronics in Agriculture*. 2008. Vol. 64. P. 72-77.
10. S. Butterworth. On the Theory of Filter Amplifiers // *Wireless Engineer*, vol. 7, 1930, pp. 536–541
11. D. Atauri Mezquida, J. Llorente Martínez, Platform for bee-hives monitoring based on sound analysis. A perpetual warehouse for swarm's daily activity // *Spanish Journal of Agricultural Research* 2009 7(4), 824-828.

12. M. Bencsik, J. Bencsik M. Baxter, Identification of the honey bee swarming process by analysing the time course of hive vibrations // *Computers and Electronics in Agriculture* 76(1):44-50, 2011. DOI: 10.1016/j.compag.2011.01.004
13. A. Robles-Guerrero, T. Saucedo-Anaya, E. Gonzalez-Ramirez, Frequency Analysis of Honey Bee Buzz for Automatic Recognition of Health Status: A Preliminary Study // *Research in Computing Science* 142 pp. 89–98; 2017.
14. В.А. Тобоев, М.С. Толстов, Ознаковий опис акустичних шумів на основі статистичної близькості функцій узагальненого середнього // *Вісник Чувашського університету* № 3. Ст. 19-30, 2010.
15. Тобоев В.А. Акустична динаміка бджолиних сімей що рояться // *Аграрна Росія*. 2010. №2. Ст. 25-28.
16. A. Qandour, I. Ahmad, D. Habibi, M. Leppard, Remote beehive monitoring using acoustic signals // *Acoustics Australia*, 42(3), p. 204-209. 2014.
17. Сергієнко А. Б. Цифрова обробка сигналів: навч. посібник 3-тє видання // СПб.: БХВ-Петербург, 2011.
18. Kaiser, James F.; Schafer, Ronald W. On the use of the I_0 -sinh window for spectrum analysis. // *IEEE Transactions on Acoustics, Speech, and Signal Processing*. 28: 105–107. (1980). doi:10.1109/TASSP.1980.1163349.
19. Сисоева С. МЭМС-микрофони // *Компоненти и технології*. 2010. № 6.
20. АХЧ та технічні характеристики ненаправленого вимірювального мікрофона Superlux-ECM999 [Електронний ресурс]
<https://superlux.su/product/superlux-ecm999/>
21. Rabiner L. R., Gold B. Theory and Application of Digital Signal Processing / Rabiner L. R., Gold B. Paramus // NJ: Prentice-Hall, 1986.
22. Продеус А.М. Про оцінки інформативності діагностичних ознак. - Праці Міжнародної дистанційної науково-практичної конференції «Сучасні інформаційні технології в діагностичних дослідженнях» // Дніпропетровськ, 2002.